Al for Material Science Talk: "Data Augmentation for Graph Regression"

Abstract:

Graph Regression is widely seen in material discovery as to predict numerical properties of molecules and polymers. In graph regression research, training sets usually have only hundreds of labeled examples. Also, the label distribution is imbalanced. Though a great number of unlabeled examples exist, they were collected from various domains and thus would be difficult to make a positive impact on target label predictions. In machine learning, data augmentation is known as the techniques that increase the amount of data by adding slightly modified copies of existing data or newly created synthetic data from existing data. It is simple yet effective. In this talk, I will introduce three graph data augmentation techniques, designed for supervised learning, imbalanced learning, and transfer learning on graph regression tasks. These studies were accepted to KDD'22, KDD'23, and NeurIPS'23.

Bio:

Meng Jiang is an Associate Professor in the Department of Computer Science and Engineering at the University of Notre Dame. He received B.E. and PhD from Tsinghua University. He was a visiting scholar at CMU and a postdoc at UIUC. He is interested in data mining, machine learning, and natural language processing. His data science research focuses on graph and text data for applications such as molecule/polymer property prediction and inverse design. He received the CAREER Award from the National Science Foundation. He has delivered 14 conference tutorials and organized seven workshops. He is a Senior Member of ACM and IEEE.

