Graph Neural Networks based Active Learning for Guided Exploration of High-yield Reactions

GIMADIEV Timur, Hokkaido University

In the realm of synthetic chemistry, the search for efficient and high-yielding reactions is essential. Traditional experimental approaches are time-consuming, resource-intensive, and often rely on intuition and trial-and-error. However, the advent of large, well-curated datasets and advances in machine learning, particularly GNNs, have opened new opportunities for accelerating reaction optimization.

Our study addresses this challenge by creating a virtual environment that mimics the process of data acquisition in a laboratory setting. By leveraging GNNs trained on the existing dataset of Suzuki-Miyaura and Buchwald-Hartwig reactions, we develop predictive models that can rapidly assess the feasibility of specific reaction conditions. Through active learning, the GNNs iteratively suggest experiments that are most likely to yield high-quality results, effectively guiding the experimentalist towards efficient reaction conditions.

The integration of GNNs and active learning not only accelerates the discovery process but also optimizes resource utilization by minimizing unnecessary experimentation. We demonstrate the potential of this approach through extensive simulations and highlight its promising implications for the field of synthetic chemistry.