**[BIOINFO 2024]**

# Condition Aware Relational Learning for Chemical Reaction Yields Prediction

Yijingxiu Lu[1], Yinhua Piao[1] and Sun Kim[1,2,*]

[1]*Department of Computer Science, Seoul National University*
[2]*AIGENDRUG Co. Ltd*
**Corresponding author: sunkim.bioinfo@sun.ac.kr*

Accurate prediction of chemical reaction yields is essential in fields such as synthetic chemistry, drug development, and materials design, as it enables more efficient experimental planning and resource allocation. However, the yield of a chemical reaction is highly sensitive to even slight variations in experimental conditions, making prediction challenging. Most existing methods overlook the exploration of the roles that different compounds play within a reaction, including how reactants, reagents, and products interact. The intrinsic characteristics of reactions often provide valuable insights for yield prediction; for instance, certain substructures of reactants undergo transformation into target products through complex reaction sequences, while condition molecules can significantly influence the final yield through intricate interactions with reactants. In this paper, we introduce a novel condition-aware relational learning model (CARL) for predicting chemical reaction yields. CARL systematically captures the complex interactions among various entities involved in a reaction—namely, reactants, products, and condition molecules—using a relational learning framework. Additionally, we develop two novel loss functions to further enhance CARL's accuracy in modeling reactions. Experimental results on benchmark datasets demonstrate that CARL achieves state-of-the-art performance using only molecular graphs as input, without requiring additional supplementary information. This approach offers a robust and interpretable method for yield prediction. Moreover, visualizations of the atom-wise interaction embeddings confirm that CARL effectively identifies crucial substructures within both reactants and condition molecules, improving its ability to capture key features that impact chemical yield prediction.