Chemistry-aware GNN for learning molecular mixture properties

Hengrui Zhang,[1] Tianxing Lai,[2] Jie Chen,[1] Arumugam Manthiram,[2] James M. Rondinelli,[1] Wei Chen[1]

Northwestern University

TEXAS The University of Texas at Austin

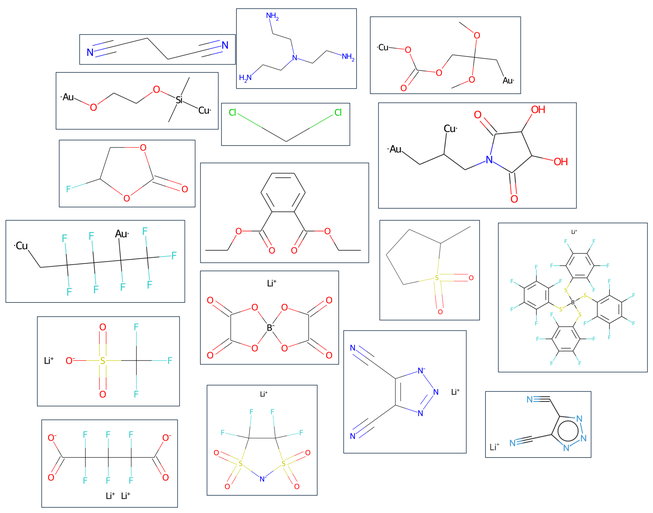
Overview

Objective: Accurate and robust prediction of molecular mixture properties

Method: MolSets = graph neural networks + permutation invariant architecture

Application: Accelerate virtual screening for, e.g., Li battery electrolyte design.

Motivation



🎯 Molecular mixtures: a vast space for materials discovery

Coolants Fuels Electrolytes

🚩 Combinatorial explosion

Select 3 from 200 → 1 million candidates!

Weight fractions → more complexity

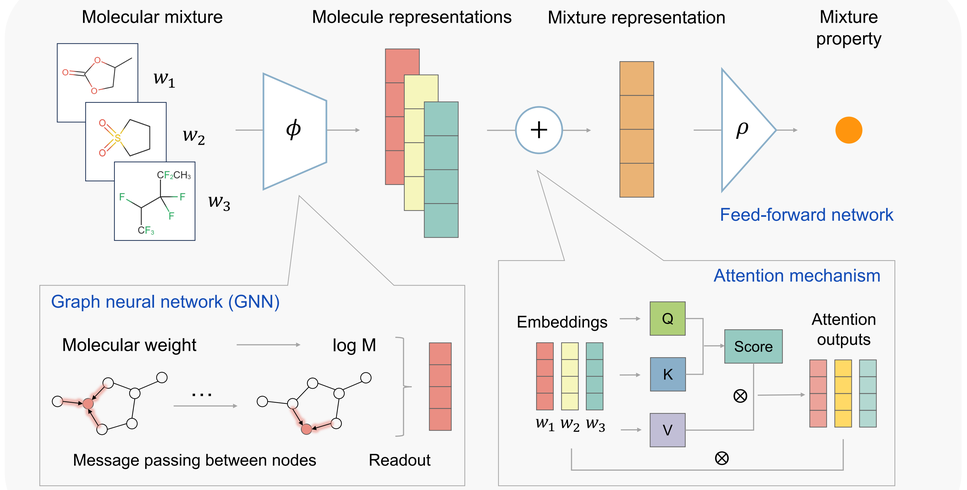
🤔 Could AI come to help?

The MolSets Model

✨ Machine learning: representation needs to capture complex “chemistry” at multiple levels

⚗️ “Chemistry” = formulation + interaction

- Molecule = atoms + geometry, bonding
- Mixture = constituents + molecular interactions
- Permutation invariance: (a, b, c) = (b, c, a)



Key ingredients

GNN to learn from molecular graphs

Deep sets to preserve permutation invariance

$$f(X) = \rho(\oplus_{x \in X}\{\phi(x)\})$$

Attention mechanism to capture interactions

Outlook

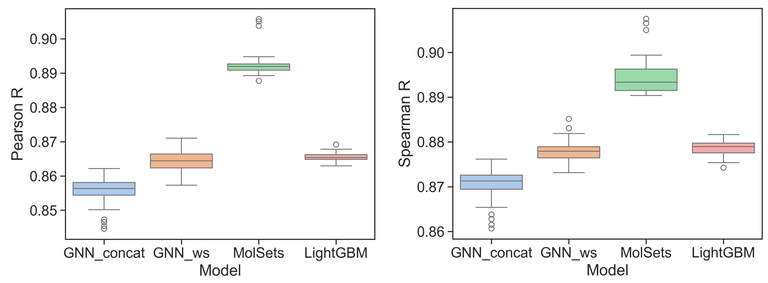
- Need more data! High-throughput experiments can help.
- Integrate to future autonomous workflows.
- Platform for efficient, (reasonably) accurate prediction, like AlphaFold 🧬 and matterverse ⚛️

Results

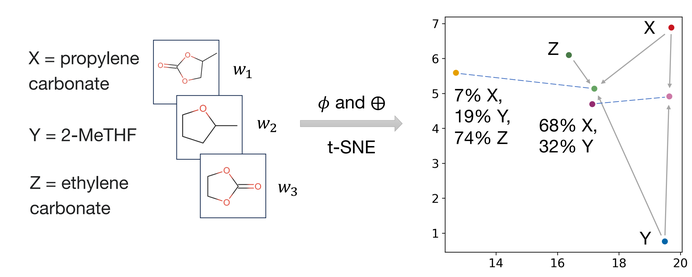
📝 Tested on a dataset of 1076 unique mixtures

Composition → conductivity for Li battery electrolytes

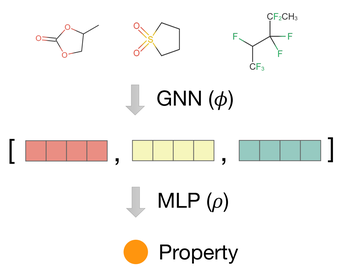
1. Benchmark and ablation tests

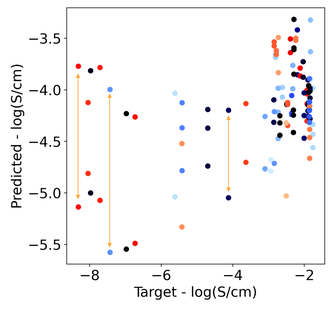


2. How is mixture representation different from linear combination of its constituents?



3. What if a model doesn’t preserve permutation invariance?





Paper: HZ, TL, JC, AM, JMR, & WC. Learning molecular mixture properties using chemistry-aware graph neural network. PRX Energy 3, 023006 (2024).

Artificial Intelligence for Materials Science Workshop

Supported by the National Science Foundation under Awards No. 2324173 and 2219489; U.S. Department of Energy, Office of Basic Energy Sciences under Award No. DE-SC0005397; and the Ryan Graduate Fellowship.

NIST