



# Evaluating Extrapolation Ability of LLM In Chemical Domain

ACL 2024  
Bangkok, Thailand

Language  
+  
Molecules

Taehun Cha and Donghun Lee  
Department of Mathematics, Korea University

## Motivation

- Human is a good extrapolator
- DNN is a good interpolator but poor at extrapolation
- **Can LLM perform extrapolation in chemical domain?**

## Experimental Setting

- LLM predict three DMA properties\* of epoxy product when
  - Adding a new resin ( $B_i$ )
  - Replacing original resin with a new resin  $B_2$
- Feed LLM SMILES, description, and added ratio of Resin  $B_i$ s
- Use GPT-4 and 917 data points from lab experiment

\*Glass transition temperature ( $T_g$ ), tan delta peak ( $\delta$ ), cross-link density ( $v_c$ )

## Results on Prediction Accuracy

	Add $B_1$		Add $B_2$		Add $B_3$		Replace $B_2$	
	$T_g$	$v_c$	$T_g$	$v_c$	$T_g$	$v_c$	$T_g$	$v_c$
Lin.Reg.	4.61	3.47	4.31	<b>2.25</b>	8.42	3.29	12.87	10.94
Ridge	<b>4.58</b>	3.49	<b>4.20</b>	2.28	8.42	3.36	12.78	10.97
Ran.For.	5.70	3.10	4.99	3.11	9.79	3.01	<b>12.62</b>	7.18
XGB	5.61	3.15	4.60	2.37	9.00	3.04	13.35	7.79
Ours	7.32	<b>2.99</b>	5.62	2.88	<b>6.40</b>	<b>2.51</b>	21.17	<b>4.67</b>

- MAE on prediction
- Omit  $\delta$  which shows a similar trend with  $T_g$

• **Not Good enough :(**

## Example Answer

LLM: "The incorporation of CTBN into an epoxy resin generally results in a **decrease in  $T_g$**  because the CTBN phase is softer and more flexible (...) The SMILES of Resin B indicates the presence of butadiene and acrylonitrile groups, which **contribute to the elastomeric properties of the resin.**" -> LLM Using its chemical knowledge!

## Results on Prediction Direction

		Average		Freq.	
		Train	Test	↑	↓
Add $B_1$	$T_g$	161	↓ 158	92	<b>358</b>
	$\delta$	0.68	↑ 0.72	<b>456</b>	322
Add $B_2$	$T_g$	161	↓ 156	91	<b>248</b>
	$\delta$	0.68	↑ 0.74	<b>271</b>	151
Add $B_3$	$T_g$	161	↓ 151	91	<b>542</b>
	$\delta$	0.68	↑ 0.74	<b>350</b>	527
Rep. $B_2$	$T_g$	161	↓ 146	18	<b>151</b>
	$\delta$	0.68	↑ 0.71	<b>176</b>	106

- Average: material properties in train and test sets.
- Freq.: Word counts on extrapolation direction
- **LLM predict a correct extrapolation direction with incorrect degree**

## 1-shot Correction

- Ask LLM to predict again with one previous prediction and ground truth value

		0-shot	1-shot
Rep. $B_2$	$T_g$	21.1	<b>12.1</b>
	$\delta$	1.11	<b>1.08</b>
	$v_c$	4.67	<b>3.89</b>

- **LLM can adjust the degree**

## Conclusion & Future Works

- **LLM uses its chemical knowledge to extrapolate to unseen materials**
  - Extrapolation direction is correct but degree is wrong
  - LLM can effectively adjust the degree with one shot correction
- **Applicability to Bayesian optimization under unseen environment?**