CrysGNN: Distilling pre-trained knowledge to enhance property prediction for crystalline materials.

Kishalay Das[1], Bidisha Samanta[1], Pawan Goyal[1], Seung-Cheol Lee[2], Satadeep Bhattacharjee[2], Niloy Ganguly[1,3]

[1]Indian Institute of Technology Kharagpur, India, [2] Indo Korea Science and Technology Center, Bangalore, India
[3]Leibniz University of Hannover, Germany

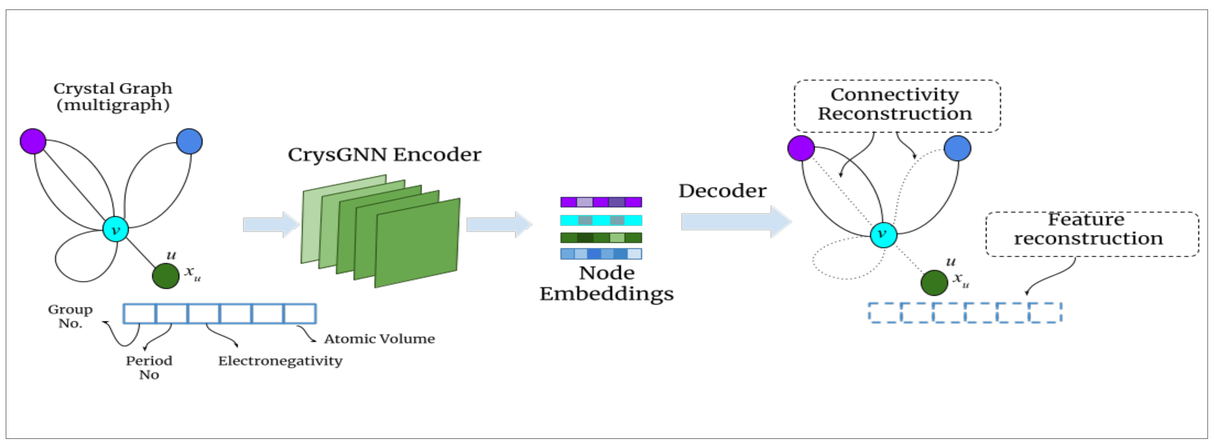
Objectives

- To build a new large **untagged crystal graph dataset** curated from popular databases of crystalline materials.
- To develop a **pre-trained GNN model** for Crystalline materials, which captures both **local (node level) chemical** and **global (graph level) structural** semantics of crystal graphs.

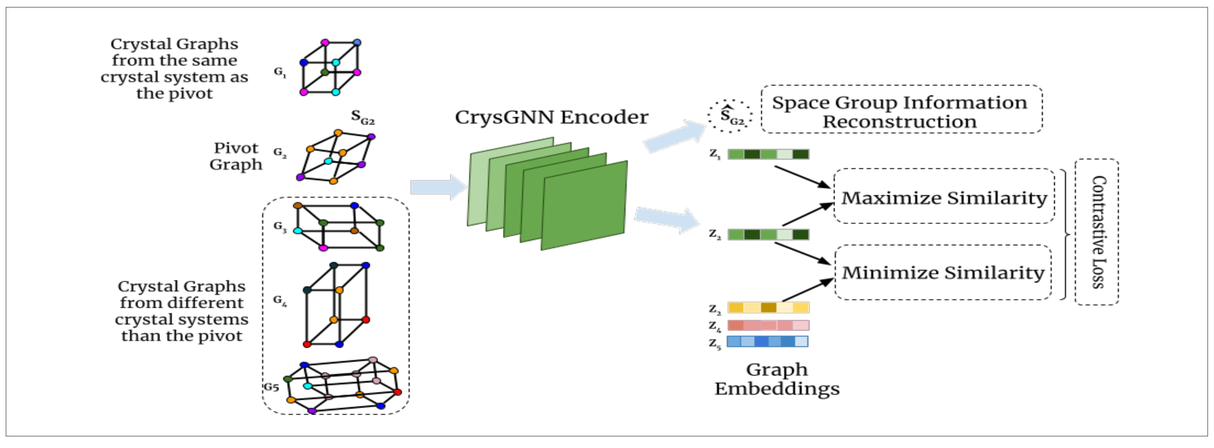
Our Contributions

- We present **CrysGNN**, a **Deep GNN Model** for Crystalline materials, which is **pre-trained** on both **node** and **graph level**.
- We curate a **new large untagged crystal dataset** with **800K crystal graphs** to pretrain CrysGNN.
- We introduce a **self-supervised graph pre-training method** which captures (a) **connectivity** of atoms, (b) **atomic properties** and (c) **graph similarity** from a large set of unlabeled crystal graph data.
- We **distill** important **structural and chemical** information of a crystal from the pre-trained CrysGNN model and pass it to the property predictor.

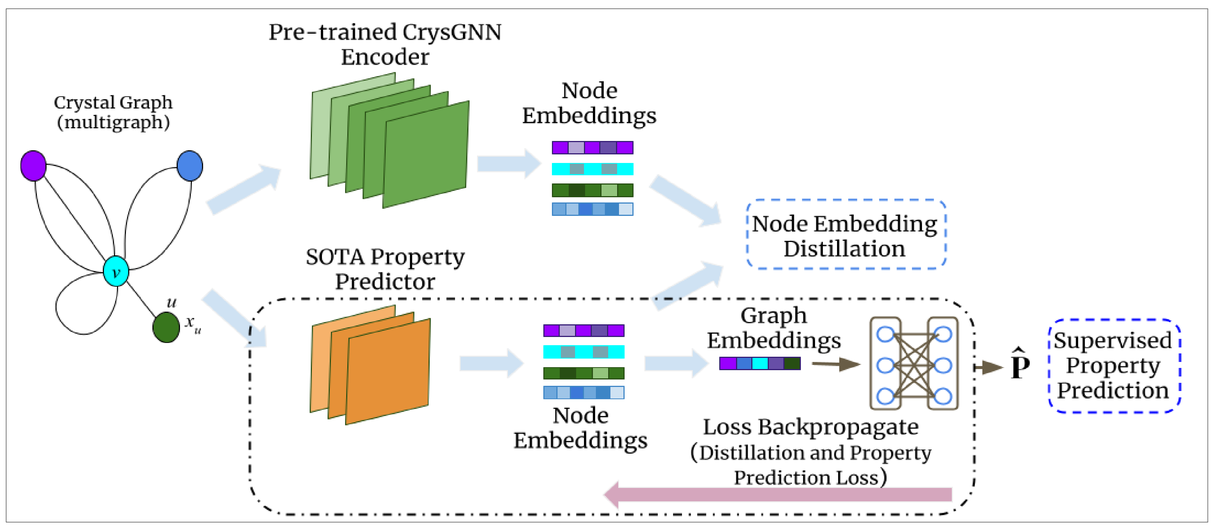
CrysGNN : Node Level Pre-training



CrysGNN : Graph Level Pre-training



Distillation and Property Prediction



Results

Property	CGCNN	CGCNN (Distilled)	CrysXPP	CrysXPP (Distilled)	GATGNN	GATGNN (Distilled)	ALIGNN	ALIGNN (Distilled)
Formation Energy	0.039	**0.032**	0.041	**0.035**	0.096	**0.091**	0.026	**0.024**
Bandgap (OPT)	0.388	**0.293**	0.347	**0.287**	0.427	**0.403**	0.271	**0.253**
Formation Energy	0.063	**0.047**	0.062	**0.048**	0.132	**0.117**	0.036	**0.035**
Bandgap (OPT)	0.200	**0.160**	0.190	**0.176**	0.275	**0.235**	0.148	**0.131**
Total Energy	0.078	**0.053**	0.072	**0.055**	0.194	**0.137**	0.039	**0.038**
Ehull	0.170	**0.121**	0.139	**0.114**	0.241	**0.203**	0.091	**0.083**
Bandgap (MBJ)	0.410	**0.340**	0.378	**0.350**	0.395	**0.386**	0.331	**0.325**
Spillage	0.386	**0.374**	0.363	**0.357**	0.350	**0.348**	0.358	**0.356**
SLME (%)	5.040	**4.790**	5.110	**4.630**	5.050	**4.950**	4.650	**4.590**
Bulk Modulus (Kv)	12.45	**12.31**	13.61	**12.70**	11.64	**11.53**	11.20	**10.99**
Shear Modulus (Gv)	11.24	**10.87**	11.20	**10.56**	10.41	**10.35**	9.860	**9.800**

Table 1:Summary of the prediction performance (MAE) of different properties in Materials Project (Top) and JARVIS-DFT (Bottom) for SOTA models, both vannila and distilled version.

Key Takeaways

- **Distilled version** of any state-of- the-art model **outperforms** the vanilla SOTA model across all the properties.
- Advantage of using distilling knowledge from a pre-trained model is we can **retrofit** it to any **downstream model** irrespective of their **architectural design**.
- Extensive experiments show its **superiority over conventional fine-tuning models** and its inherent ability to **remove DFT-induced bias**.

References

[1] Xie et. al; **CGCNN**; **Phys. Rev. Lett. 2018**
[2] Das et. al; **CrysXPP**; **npj Computational Materials 2021**
[3] Choudhary, et. al; **ALIGNN**; **npj Computational Materials 2021**

Contact Information

Contact

Paper Details

Github