AMRITA VISHWA VIDYAPEETHAM

GRAPH NEURAL NETWORK APPROACHES FOR PREDICTING MECHANICAL PROPERTIES OF CRYSTALLINE MATERIALS

Presented By : Team 6

Team members

• A Prithvi	CB.SC.U4AIE23201
• Dalli Krishan Preetham Reddy	CB.SC.U4AIE23222
• S Dhanush	CB.SC.U4AIE23243

Introduction

- Accurate prediction of material properties helps speed up the discovery and design of new materials, but traditional methods like DFT are too slow and computationally expensive.

- The study uses Graph Neural Networks (GNNs), Graph Convolutional Networks (GCNs) and GraphSAGE to predict key mechanical properties of crystalline materials — Formation energy and band gap.

- Crystal structures are modeled as graphs, where atoms are nodes (with physicochemical features) and edges represent inter-atomic interactions. The models are trained on a curated dataset and evaluated using mean absolute error (MAE) and coefficient of determination (R^2).

Problem Statement

- Predicting material properties like Formation Energy and Band Gap is essential for new material discovery.
- Traditional methods like DFT simulations are accurate but computationally expensive.
- Challenge: Finding which CBFV featurization technique and GNN architecture yield the best accuracy and interpretability.

Cleaning Dataset

- The initial dataset, Perovskite.csv, contained approximately 4600 data points representing various perovskite compounds and their material properties. However, a detailed inspection revealed that the dataset included numerous missing values and outliers across several critical columns.

- To ensure data reliability and model stability, a comprehensive preprocessing pipeline was applied, which involved:
  - Removal of incomplete records with missing or undefined feature values.
  - Outlier detection and elimination using statistical thresholds to maintain consistency in physical property ranges.
  - Feature normalization and encoding for uniform scaling across descriptors.

- After this process, the refined dataset named Perovskite_data_cleaned.csv was obtained, containing just over 1,000 high-quality data points.

Compostion Based Feature Vectors

- CBFV encodes materials as composition-derived descriptors. Features include atomic, electronic, thermodynamic, and structural categories.

- Example categories:
  - Atomic size: Atomic weight, radius variations.
  - Electronic: Electronegativity, ionization energy.
  - Thermodynamic: Melting point, heat of fusion.

- The featurizer selects relevant subsets from CBFV presets such as Oliynyk or Jarvis. Out of which Oliynyk gave the best results.

Graph Representation of Materials

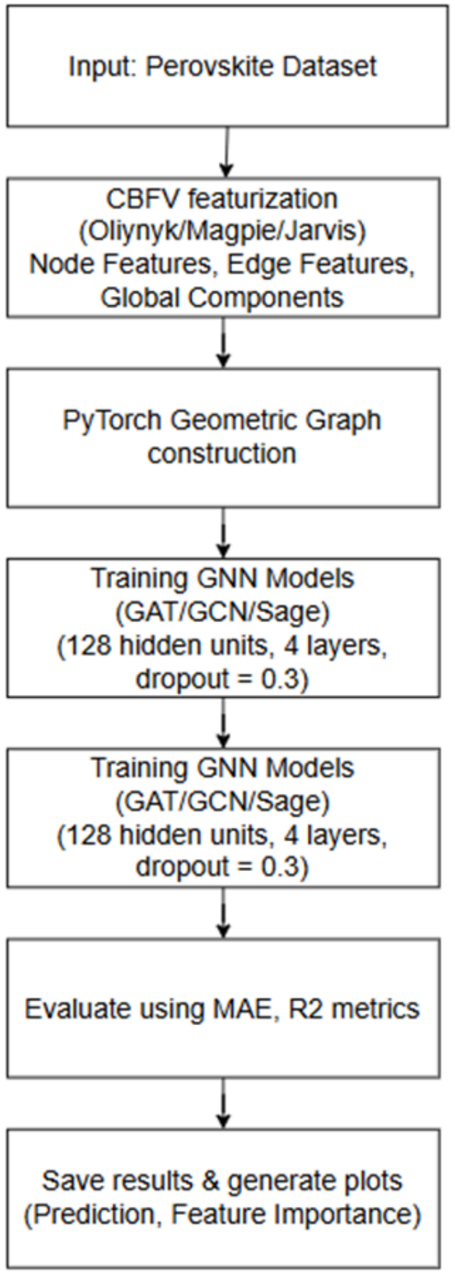
- Graph Representation: Each material structure from Perovskite_data_cleaned.csv is converted into a heterogeneous graph, where atoms are represented as nodes and interatomic bonds or spatial interactions are represented as edges.

- Feature Encoding: Each node carries CBFV-based atomic descriptors, while edges encode pairwise atomic or structural relationships. Additionally, global crystal-level features (u) capture composition-based statistics for the entire material.

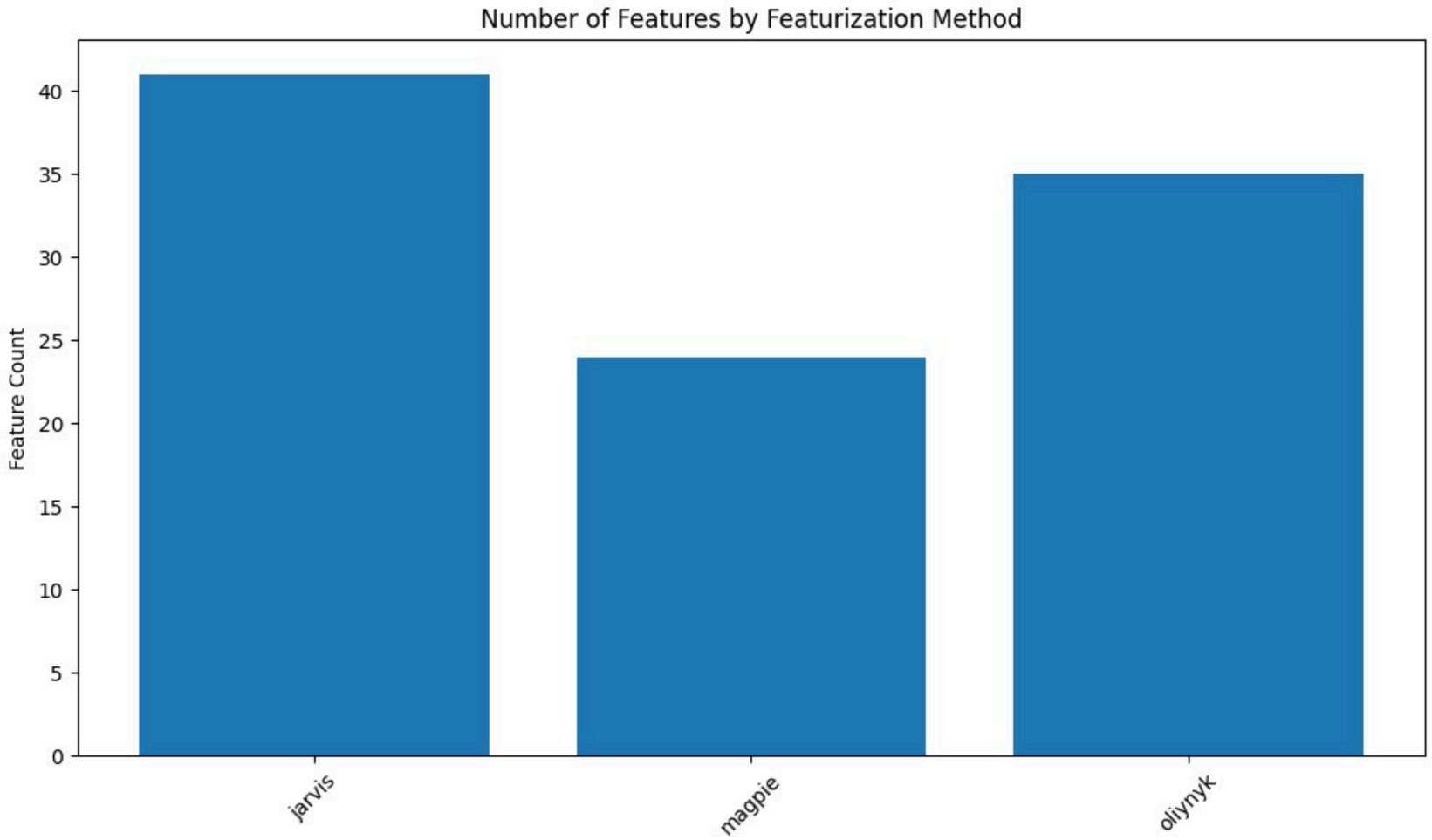
- Implementation Framework: All graphs are constructed and processed using the PyTorch Geometric (PyG) framework, enabling efficient graph-based learning and seamless integration with GAT, GCN, and GraphSAGE architectures.

Methodology

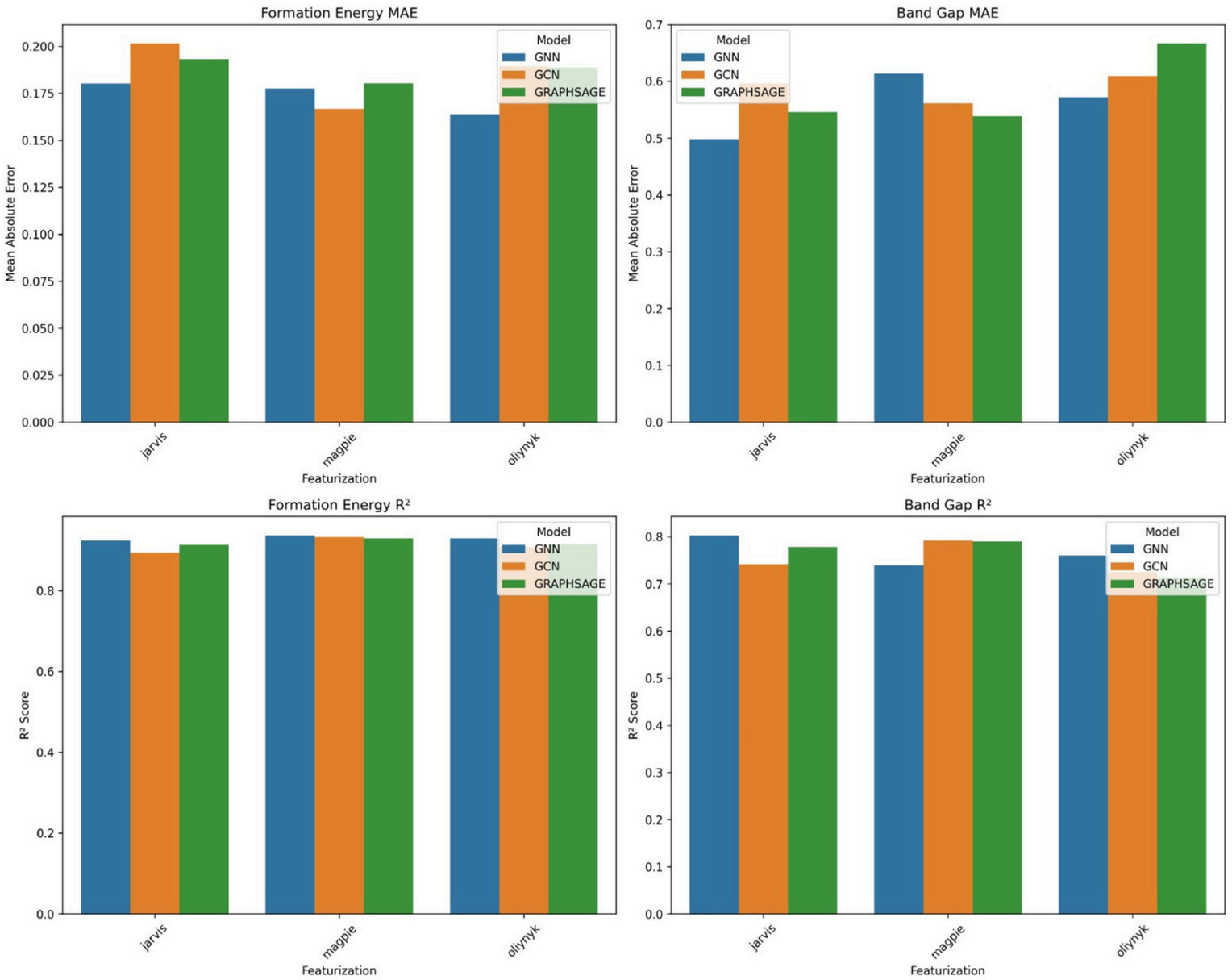
Goal: The goal of the project is to build and compare GAT, GCN, and GraphSAGE models on different featurizations — Jarvis, Magpie, and Oliynyk.



Feature Comparison:

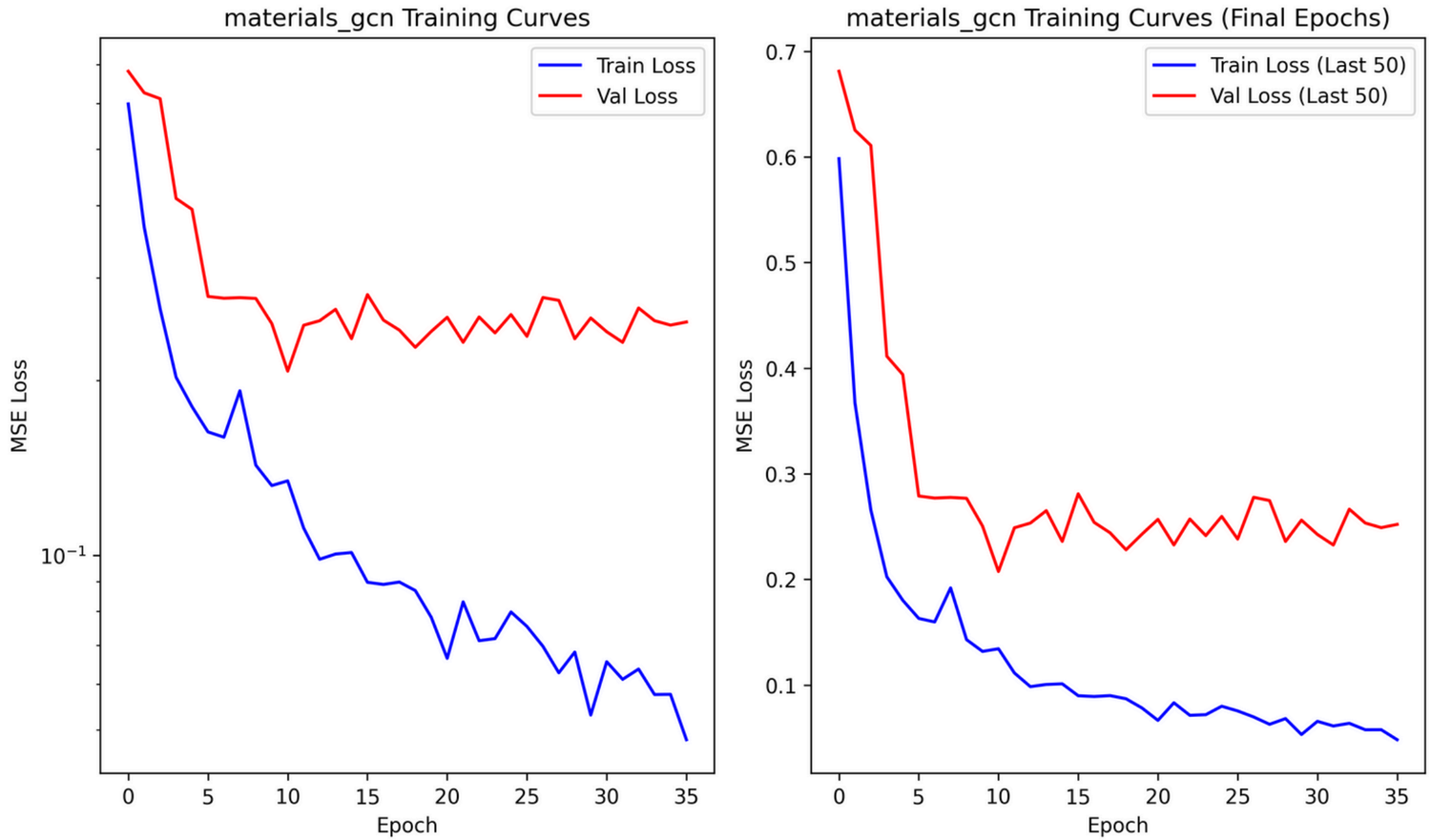


Feature comparison:



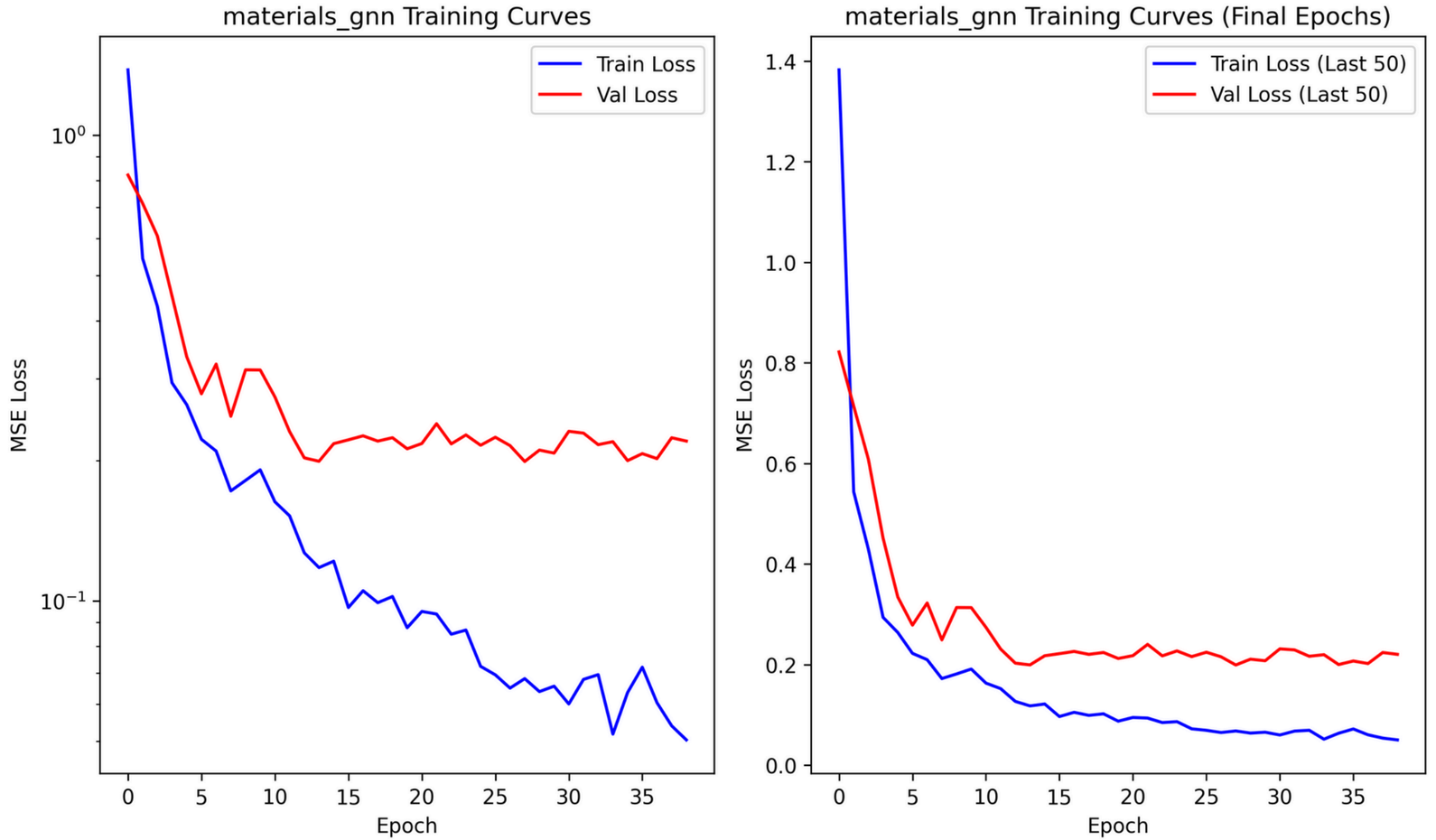
Results

GCN Training Curves



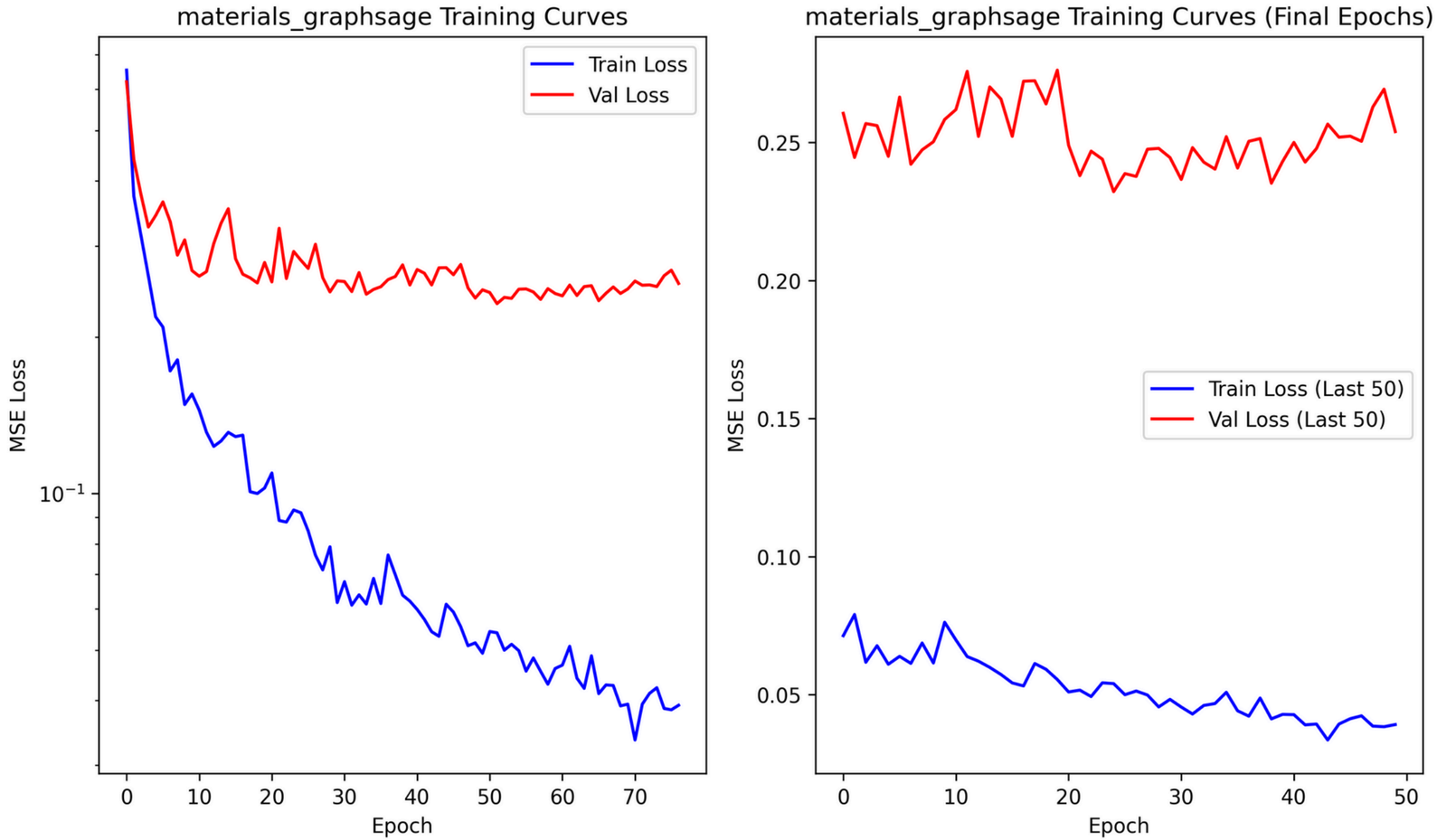
Results

GNN Training Curves



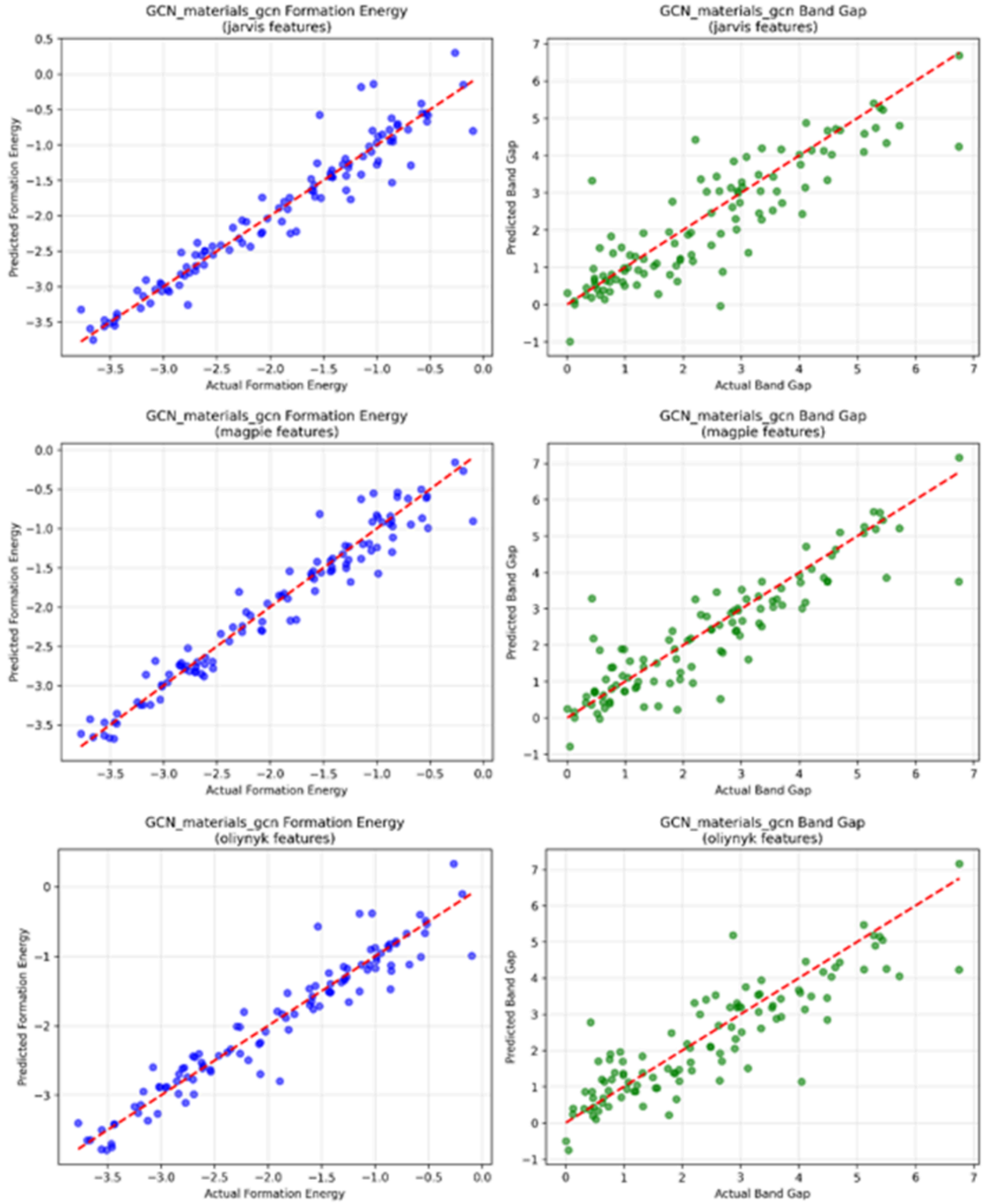
Results

GraphSAGE Training Curves



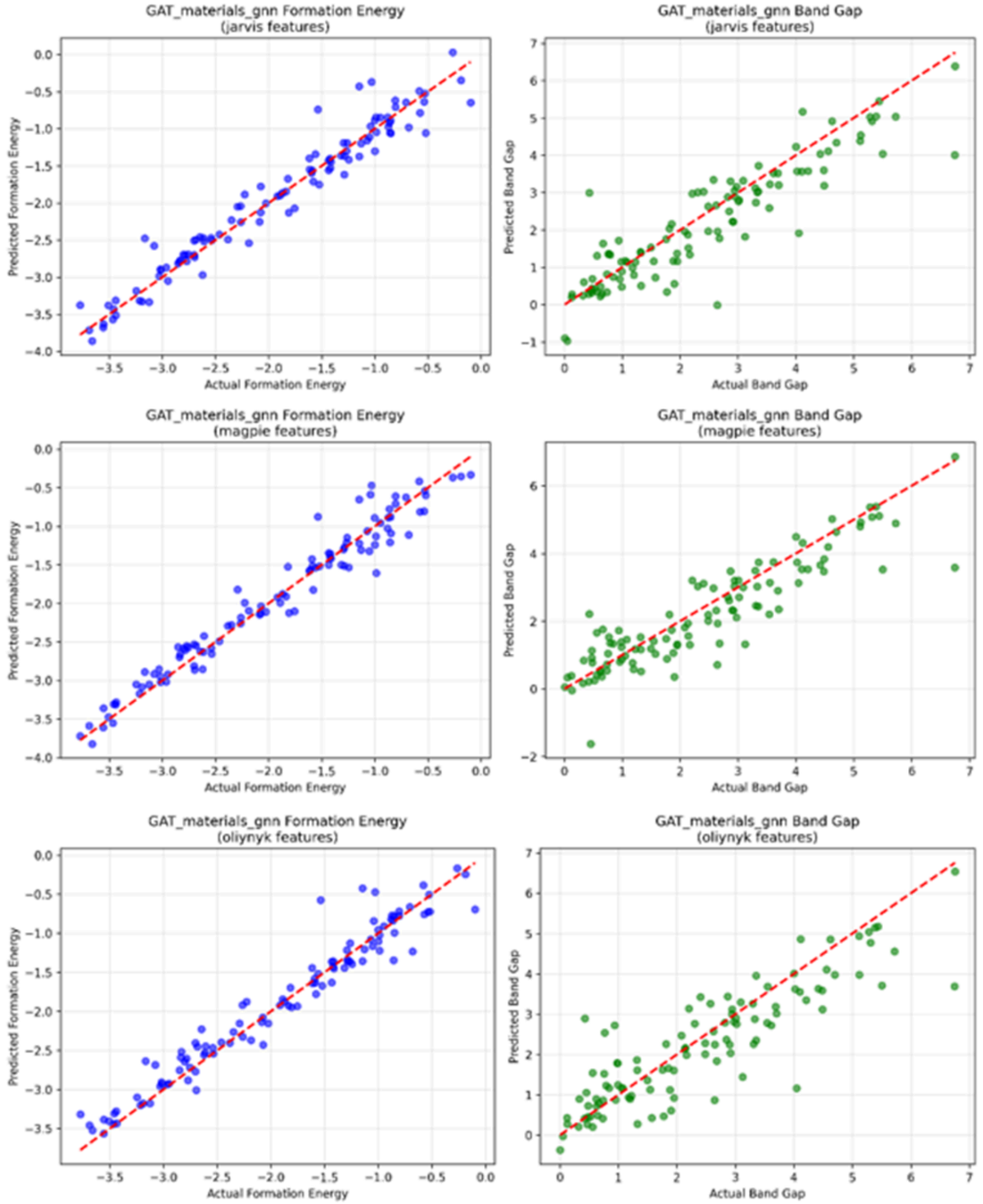
Results

GCN Predictions
(using Jarvis, Magpie, Oliynyk)



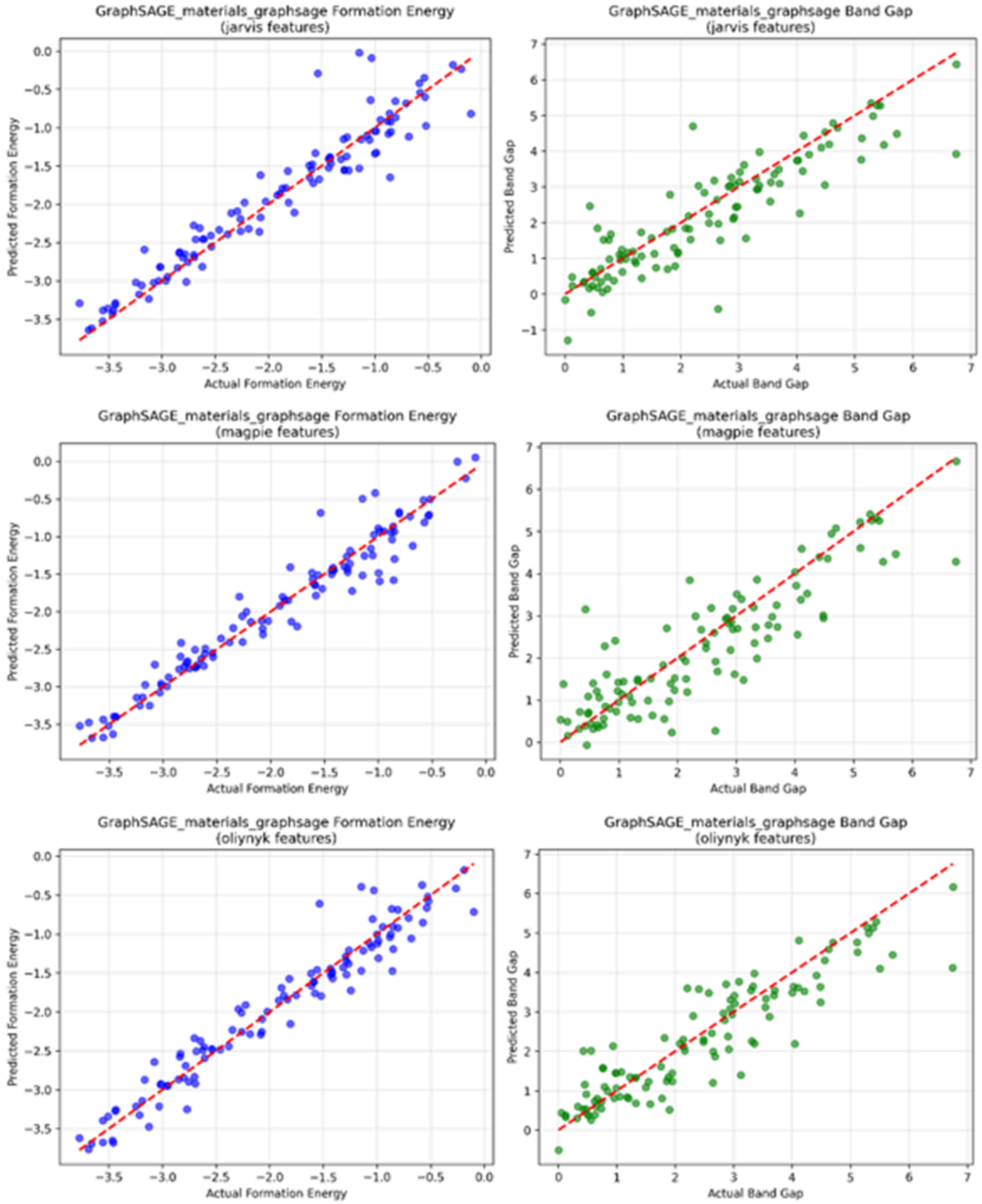
Results

GAT Predictions
(using Jarvis, Magpie, Oliynyk)



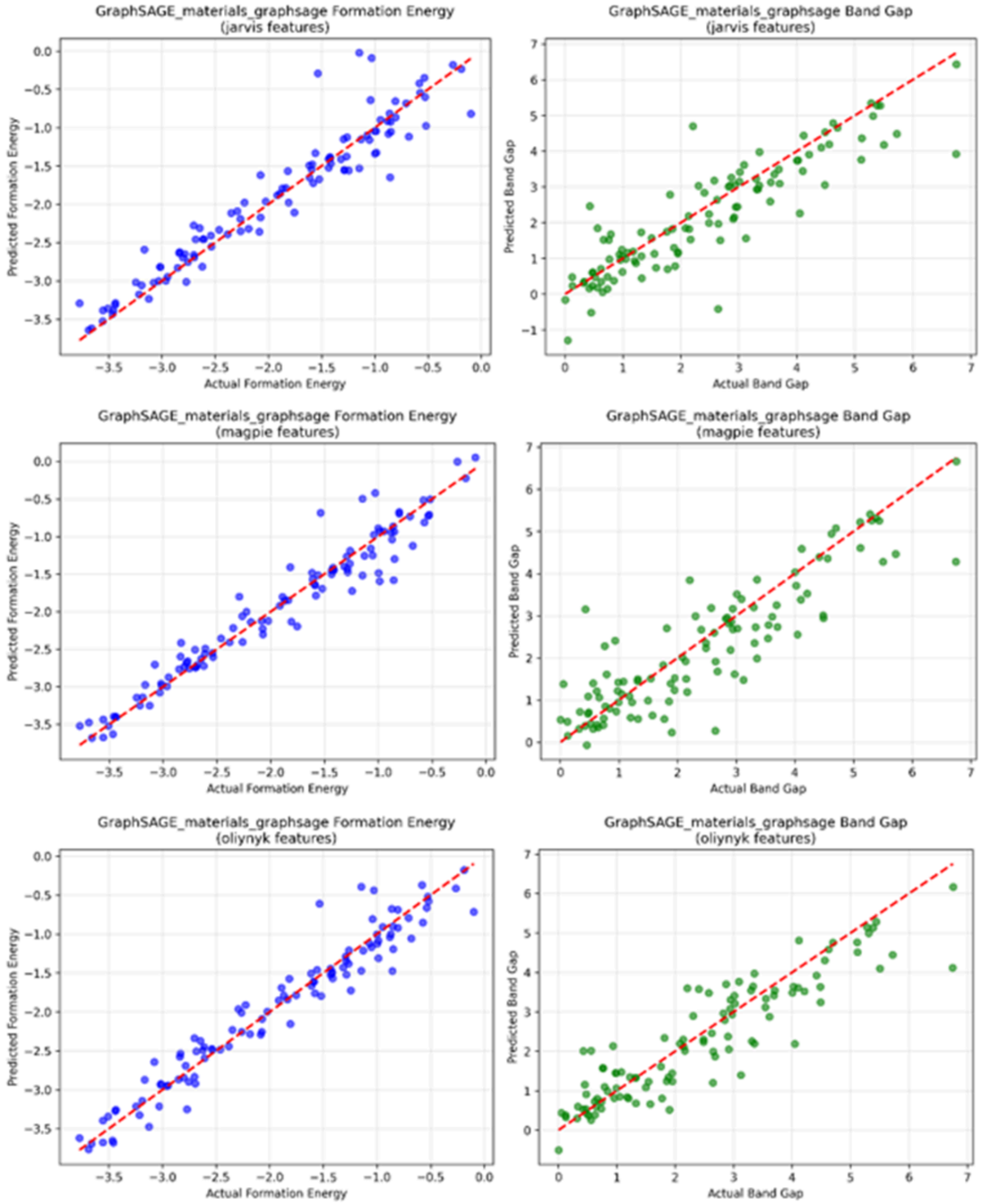
Results

GraphSAGE Predictions
(using Jarvis, Magpie, Oliynyk)



Results

GraphSAGE Predictions
(using Jarvis, Magpie, Oliynyk)



Results

Metrics for materials_gcn (using Jarvis, Magpie, Oliynyk)

model	split	featurizer	formation_mae	formation_mse	formation_rmse	formation_r2	bandgap_mae	bandgap_mse	bandgap_rmse	bandgap_r2
materials_gcn	train	-	0.05728	0.00536	0.073211	0.993916	0.153766	0.038079	0.195138	0.982695
materials_gcn	test	-	0.167923	0.06887	0.262432	0.925574	0.543767	0.550775	0.742142	0.794443
materials_gcn	train	jarvis	0.068081	0.007572	0.08702	0.991405	0.184946	0.058012	0.240858	0.973636
materials_gcn	test	jarvis	0.164781	0.06663	0.258128	0.927995	0.580382	0.661698	0.813448	0.753045
materials_gcn	train	magpie	0.064087	0.006915	0.083159	0.992151	0.192652	0.071141	0.266723	0.96767
materials_gcn	test	magpie	0.155361	0.047729	0.218469	0.948421	0.510056	0.551338	0.742522	0.794233
materials_gcn	train	oliynyk	0.074553	0.009939	0.099695	0.988719	0.217391	0.096089	0.309982	0.956332
materials_gcn	test	oliynyk	0.185529	0.074808	0.27351	0.919157	0.60123	0.64935	0.805823	0.757654

Results

Metrics for materials_gnn (using Jarvis, Magpie, Oliynyk)

model	split	featurizer	formation_mae	formation_mse	formation_rmse	formation_r2	bandgap_mae	bandgap_mse	bandgap_rmse	bandgap_r2
materials_gnn	train	-	0.043173	0.002893	0.053788	0.996716	0.142564	0.036842	0.191942	0.983257
materials_gnn	test	-	0.152792	0.049004	0.221368	0.947043	0.543933	0.620945	0.788001	0.768255
materials_gnn	train	jarvis	0.05321	0.004758	0.06898	0.994599	0.145966	0.040201	0.200503	0.98173
materials_gnn	test	jarvis	0.15525	0.049617	0.222749	0.94638	0.518627	0.541781	0.736058	0.7978
materials_gnn	train	magpie	0.065193	0.006535	0.080838	0.992583	0.182402	0.059298	0.243512	0.973052
materials_gnn	test	magpie	0.148291	0.039915	0.199788	0.956864	0.549994	0.568837	0.754213	0.787702
materials_gnn	train	oliynyk	0.057993	0.005299	0.072797	0.993985	0.184323	0.065654	0.25623	0.970163
materials_gnn	test	oliynyk	0.15297	0.050196	0.224045	0.945754	0.581531	0.673108	0.820431	0.748787

Results

Metrics for materials_graphsage (using Jarvis, Magpie, Oliynyk)

model	split	featurizer	formation_mae	formation_mse	formation_rmse	formation_r2	bandgap_mae	bandgap_mse	bandgap_rmse	bandgap_r2
materials_graphsage	train	-	0.04961	0.003772	0.061414	0.995719	0.145218	0.037977	0.194878	0.982741
materials_graphsage	test	-	0.164534	0.06694	0.258728	0.92766	0.501388	0.484355	0.695956	0.819232
materials_graphsage	train	jarvis	0.055228	0.004929	0.070208	0.994405	0.154495	0.044564	0.211102	0.979748
materials_graphsage	test	jarvis	0.184654	0.080526	0.283771	0.912977	0.543754	0.617742	0.785966	0.76945
materials_graphsage	train	magpie	0.061082	0.006404	0.080023	0.992732	0.182378	0.066764	0.258388	0.969659
materials_graphsage	test	magpie	0.155014	0.052939	0.230085	0.94279	0.576755	0.644909	0.803062	0.759311
materials_graphsage	train	oliynyk	0.064399	0.00642	0.080122	0.992714	0.176023	0.067893	0.260563	0.969146
materials_graphsage	test	oliynyk	0.169575	0.055906	0.236445	0.939584	0.515897	0.489477	0.699626	0.817321

THANK YOU

Presented By : Team 6