Graph Neural Network Approaches for Predicting Mechanical Properties of Crystalline Materials

A Prithvi
Department of Artificial Intelligence
Amrita Vishwa Vidyapeetham
Coimbatore, India
cb.sc.u4aie23201@cb.students.amrita.edu

Dalli Krishan Preetham Reddy
Department of Artificial Intelligence
Amrita Vishwa Vidyapeetham,
Coimbatore, India
cb.sc.u4aie23222@cb.students.amrita.edu

S Dhanush
Department of Artificial Intelligence
Amrita Vishwa Vidyapeetham
Coimbatore, India
cb.sc.u4aie23243@cb.students.amrita.edu

Dr. Kritesh Kumar Gupta
Department of Artificial Intelligence
Amrita Vishwa Vidyapeetham
Coimbatore, India
g_kriteshkumar@cb.students.amrita.edu

Abstract:

Prediction of material properties is an important aspect of material science, where accuracy, design and discovery of novel materials are crucial. However, usage of conventional computational techniques of such as Density Functional Theory (DFT) on feasible materials is limited by computational expenses and time constraints. This makes the task of efficiently estimating critical properties of materials arduous because of the intricate atomic interactions and non- linear relationships among structural features, which directly influence material stability and properties.

In order to predict essential mechanical properties, we apply Graph Neural Network (GNN) based architectures, specifically Graph Convolutional Network (GCN), Graph Attention Network (GAT) and GraphSAGE. These GNNs can predict key electronic and mechanical properties by learning interatomic relationships after modelling each atom as a node and each bond as an interaction between atoms to create meaningful graphs, thus increasing both speed and accuracy of predictions without heavy computational loads.

We have found that all GNN architectures outperform conventional machine learning methods in both Mean Absolute Error (MAE) and R square values. Among the models, GraphSAGE achieved the best performance by showing high accuracy in prediction of the formation energy and band gap.

These results show that GNNs can effectively capture complex atomic relationships within crystalline structures, offering a scalable, accurate and efficient approach for materials property prediction, paving the way for rapid identification of materials with desired properties.

1. Introduction

    a. Motivation of the study:

    The increasing demand for materials with custom mechanical and electronic properties is increasing rapidly in technological fields such as semiconductors and aerospace. Traditionally, these kinds of materials will be found using Density Functional Theory (DFT) or experimental simulation. Though they are accurate, the process is computationally intensive, making them impractical in case of screening thousands of candidate materials. At the same time, they often fail to understand

atomic connectivity and crystal geometry. With the advancement in the domain of artificial intelligence, especially Graph Neural Networks overcome these limitations by representing the materials as graphs, thus effectively preserving their atomic structures and relationships. This enables the models to learn and understand how local and global arrangements of atoms influence the material properties, giving it a boost over traditional models for accurate predictions.

b. Problem statement

Prediction of material properties such as formation energy and band gap is a complex challenge due to the complex atomic interactions and crystal properties. Traditional approaches like DFT rely on handcrafted descriptors which yield accurate results, but are computationally expensive. Simpler regression models struggle to capture the nonlinear relationships between atoms. Moreover, raw material datasets often contain missing or noisy data, which makes the model training unreliable. Therefore, this is a critical need for a data-efficient, structure-aware predictive model that can truly understand the underlying principles and generalize across diverse crystal systems. This project addresses the problem by applying three advanced GNN architectures, namely GCN, GAT and GraphSage to represent the crystal structures as graph and create meaningful representations and predict formation energy and bang gap with high accuracy.

c. Plan of action

The project begins by collecting a crystalline materials dataset containing material property information. The initial dataset is then pre-processed by removing incomplete and outlier data, resulting in a refined dataset. Each material is then converted to a graph structure where the atoms are the nodes contains all chemical properties such as atomic number and electronegativity, and the atomic bonds are the edges.

Three Graph Neural Network architectures i.e. Graph Convolutional Network (GCN), Graph Attention Network (GAT) and GraphSAGE, were implemented and made to learn the local and global structural patterns. The model will predict formation energy and band gap, two important features of materials. These models use aggregation techniques and message passing to learn atomic relationships efficiently. The performance of the models was evaluated using Mean Absolute Error (MAE) and Coefficient of Determination (R-square) metrics on both training and testing sets.

Finally, an analysis of comparison between all three models will be done to see which model can generalize properly and apply the graphing techniques properly to make accurate predictions. The study concludes that GNNs are a scalable and efficient method for material property prediction, and will help in accelerating the process of material discovery.

2. Literature Review

S.No.	Referred paper	Summary of the work	Relevance to your problem statement	Research Gap
1	Xie, T. & Grossman, J.C. (2018). Crystal Graph Convolutional Neural Networks for an Accurate and Interpretable Prediction of Material Properties. *Physical Review Letters*, 120, 145301. DOI: 10.1103/PhysRevLett.120.145300	Developed CGCNN framework that directly learns material properties from crystal structure using graph representation. Achieved MAE of 0.039 eV/atom for formation energy and 0.388 eV for band gap prediction on Materials Project dataset.	Directly addresses the core problem of predicting formation energy and band gap using graph-based representations, establishing baseline GCN architecture for crystalline materials	Limited to GCN architecture only; does not explore attention mechanisms or alternative aggregation strategies; uses only distance-based edge features.
2	Chen, C., Ye, W., Zuo, Y., Zheng, C. & Ong, S.P. (2019). Graph Networks as a Universal Machine Learning Framework for Molecules and Crystals. Chemistry of Materials, 31(9), 3564-3572. DOI: 10.1021/acs.chemmater.9b01294	Introduced MEGNet (Materials Graph Network) with edge, node, and global state attributes. Outperformed prior ML models with training on ~60,000 crystals from Materials Project for formation energy, band gaps, and elastic moduli prediction.	Demonstrates universality of graph networks for materials prediction; provides comprehensive framework including global features which aligns with the project's multi-level feature approach.	Does not specifically compare GAT and GraphSAGE architectures; computational complexity increases with large graphs; limited analysis of feature importance.

3	Cheng, J., Zhang, C. & Dong, L. (2021). A geometric-information-enhanced crystal graph network for predicting properties of materials. Communications Materials, 2, 92. DOI: 10.1038/s43246 021-00194-3	Proposed GeoCGNN incorporating complete geometric information using distance vectors and mixed basis functions (Gaussian radial basis + plane waves). Achieved 25.6% improvement over CGCNN and 14.3% over MEGNet for formation energy prediction.	Addresses the geometric structure limitation in basic GNNs; demonstrates importance of complete spatial information which relates to the project's comprehensive featurization approach using CBFV.	Focuses primarily on geometric encoding; does not explore different GNN architectures like GAT or GraphSAGE; computational overhead of mixed basis functions.
4	Park, C.W. & Wolverton, C. (2020). Developing an improved crystal graph convolutional neural network framework for accelerated materials discovery. Physical Review Materials, 4, 063801. DOI: 10.1103/PhysRevMaterials.4.0 63801	Developed iCGCNN with improved crystal graph representation and convolution operations. Enhanced feature aggregation and demonstrated superior performance on formation energy and band gap predictions compared to original CGCNN.	Directly relevant to improving GCN architecture for formation energy and band gap prediction; addresses some limitations of basic CGCNN framework.	Still limited to GCN-based architecture; does not compare with attention-based mechanisms; lacks exploration of different message passing strategies.
5	Reiser, P., Neubert, M., Eberhard, A., Torresi, L., Zhou,	Comprehensive review of	Provides comprehensi	Review paper - does not propose new

	C., et al. (2022). Graph neural networks for materials science and chemistry. Communications Materials, 3, 93. DOI: 10.1038/s43246 022-00315-6	GNN applications in materials science covering CGCNN, MEGNet, SchNet, and various architectures. Discussed formation energy, band gap, and synthesizability prediction methodologies.	ve overview of GNN architectures and their applications to the exact problem domain; validates the choice of using multiple GNN architectures (GCN, GAT, GraphSAGE).	methodology; identifies need for systematic comparison of different GNN architectures on same datasets.
6	Louis, S.Y., Zhao, Y., Nasiri, A., Wang, X., Song, Y., et al. (2020). Graph convolutional neural networks with global attention for improved materials property prediction. Physical Chemistry Chemical Physics, 22, 18141-18148. DOI: 10.1039/D0CP01474E	Introduced global attention mechanism to CGCNN framework, allowing model to focus on most relevant atomic interactions. Demonstrated improved prediction accuracy for formation energy and band gap on Materials Project dataset.	Directly addresses attention mechanism integration with GCN, relevant to the GAT architecture used in the project; shows benefit of attention for materials property prediction.	Limited to incorporating attention in GCN framework; does not explore standalone GAT architecture; lacks comparison with GraphSAGE aggregation methods.
7	Choudhary, K. & DeCost, B. (2021). Atomistic Line Graph Neural Network for improved materials property predictions. npj Computational Materials, 7, 185. DOI: 10.1038/s41524 021-00650-1	Developed ALIGNN (Atomistic Line Graph Neural Network) that considers both bond angles and	Addresses limitation of only using pairwise interactions; demonstrates importance of angular information which	Different graph construction paradigm (line graphs); computational complexity higher than standard GNNs; not directly comparable to

		bond lengths. Achieved state-of-the-art performance on formation energy, band gap, and other properties using JARVIS-DFT dataset	complements the project's comprehensive CBFV featurization approach.	GAT/GraphSAGE architectures.
8	Schütt, K.T., Sauceda, H.E., Kindermans, P.J., Tkatchenko, A. & Müller, K.R. (2018). SchNet – A deep learning architecture for molecules and materials. Journal of Chemical Physics, 148, 241722. DOI: 10.1063/1.5019779	Introduced SchNet using continuous filter convolutional layers for modeling quantum interactions. Achieved high accuracy on molecular and materials property prediction including formation energies.	Demonstrates continuous filter approach for materials; addresses quantum mechanical aspects of atomic interactions relevant to DFT level accuracy requirements.	Designed primarily for molecular systems; continuous filters may not capture periodic boundary conditions well; different paradigm from message passing GNNs.
9	Jha, D., Gupta, V., Wei, L., Miikkulainen, R., et al. (2024). Reinforcing crystal material property prediction with comprehensive message passing via deep graph networks. Computational Materials Science, 236, 112819. DOI: 10.1016/j.commatsci.2024.112819	Proposed CGGAT model combining GAT and Gated Graph ConvNets for comprehensive message passing between atom and edge graphs. Demonstrated improved accuracy for	Highly relevant combines GAT with enhanced message passing; addresses the need for comprehensive feature aggregation similar to project's multi-	Increased model complexity; requires more computational resources; limited analysis of individual architecture contributions (GAT vs GatedGCN).

		formation energy and band gap prediction.	architecture approach.	
10	Xu, Y., Verma, V., Naveed, H., et al. (2024). A machine learning-based crystal graph network and its application in development of functional materials. Materials Genome Engineering Advances, 2(2), e38. DOI: 10.1002/mgea.38	Comprehensive review and application study of CGCNNs for formation energy, band gap, Fermi energy, and elastic properties. Demonstrated practical applications in materials development with focus on perovskites.	Directly relevant to perovskite materials used in the project; validates application of GNN to the specific material class and properties of interest.	Review-focused with limited novel methodology; does not systematically compare different GNN architectures; lacks detailed ablation studies.
11	Yan, K., Liu, Y., Lin, Y. & Ji, S. (2021). Periodic graph transformers for crystal material property prediction. Advances in Neural Information Processing Systems, 34, 15066-15080.	Introduced periodic graph transformers that explicitly model periodicity of crystal structures. Achieved superior performance on Materials Project benchmarks for formation energy and band gap.	Addresses periodicity limitation of standard GNNs; transformer-based approach provides alternative to traditional message passing relevant for understanding architectural choices.	Transformer-based different paradigm from GCN/GAT/GraphSAGE; higher computational cost; may not scale well to large crystal structures.
12	Goodall, R.E.A. & Lee, A.A. (2020). Predicting materials properties without crystal structure: Deep representation learning from stoichiometry. Nature Communications, 11,	Developed CrabNet using composition-based features and	Provides context for composition-based features (CBFV) used	Does not use structural information fundamentally different approach; shows upper limit of

	6280. DOI: 10.1038/s41467-020-19964-7	attention mechanisms without requiring crystal structure. Demonstrated competitive performance on formation energy and band gap prediction.	in the project; validates importance of elemental property featurization complementary to structural information.	composition-only methods, highlighting value of graph-based structural encoding.
13	Omee, S.S., Louis, S.Y., Fu, N., Wei, L., et al. (2022). Scalable deeper graph neural networks for high-performance materials property prediction. Npj Computational Materials, 8, 63. DOI: 10.1038/s41524-022-00759-x	Introduced DenseGNN addressing over smoothing in deep GNNs through dense connections and attention mechanisms. Achieved state-of the-art results on formation energy, band gap, and bulk modulus prediction.	Addresses scalability and depth limitations of GNNs; demonstrates that deeper networks with proper architecture can improve prediction accuracy relevant to model design choices.	Focuses on depth rather than architectural diversity; increased parameters may lead to overfitting on smaller datasets; does not compare GAT vs GraphSAGE specifically.
14	Karamad, M., Magar, R., Shi, Y., Siahrostami, S., Gates, I.D. & Barati Farimani, A. (2020). Orbital graph convolutional neural network for material property prediction. Physical Review Materials, 4, 093801. DOI: 10.1103/PhysRevMaterials.4.093801	Developed OGCNN incorporating orbital interaction information into graph representation. Demonstrated improved accuracy for band gap and formation energy predictions by encoding	Addresses quantum mechanical aspects of atomic interactions; shows value of incorporating physical principles into features, validating project's comprehensive CBFV	Requires orbital information which may not always be available; increased computational complexity; limited to specific material types where orbital theory applies.

		quantum mechanical features.	featurization strategy.	
15	Joshi, R.P., Eickholt, J., Li, L., Fornari, M., Barone, V. & Perkins, J.E. (2021). Machine learning the voltage of electrode materials in metal-ion batteries. ACS Applied Materials & Interfaces, 11(20), 18494-18503. DOI: 10.1021/acsami.9b04933	Applied CGCNN and feature engineering for predicting formation energies of battery materials. Demonstrated importance of domain-specific featurization combined with graph networks.	Validates combination of handcrafted features (like CBFV) with graph networks; shows application to energy materials including formation energy prediction.	Focused on battery materials - specific domain; does not explore GAT or GraphSAGE architectures; limited comparison of different featurization schemes.

3. Methodology
    a. Mathematical formulation, architecture of the used ML/DL/NLP/BERT models

Problem Definition: The model predicts two continuous material properties:

- Formation energy (E_{form})
- Band gap (E_g)

Each material is represented as a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, X, E, u)$:

- $\mathcal{V}$: nodes representing atoms
- $\mathcal{E}$: edges representing atomic bonds
- $X \in \mathbb{R}^{N \times F_v}$: node feature matrix (atomic-level descriptors)
- $E \in \mathbb{R}^{M \times F_e}$: edge feature matrix (bond descriptors)
- $u \in \mathbb{R}^{F_g}$: global (composition-based) features

The task is to learn a mapping $f_\theta: \mathcal{G} \rightarrow [E_{form}, E_g]$.

$$\hat{y} = f_\theta(\mathcal{G}) = f_\theta(\mathcal{V}, \mathcal{E}, X, E, u)$$

Model Architecture

Three models were used to compare graph-based learning architectures:

1. MaterialsGNN (GAT-based)

- Uses Graph Attention Convolutions (GATConv) with multi-head attention.
- For each node i, attention is computed as:

$$e_{ij} = \text{LeakyReLU}(a^\top [Wh_i \;||\; Wh_j])$$

$$\alpha_{ij} = \frac{\exp(e_{ij})}{\sum_{k \in \mathcal{N}(i)} \exp(e_{ik})}$$

$$h_i' = \sigma\left(\sum_{j \in \mathcal{N}(i)} \alpha_{ij} W h_j\right)$$

where:

- a: attention vector
- W: learnable weight matrix
- $\mathcal{N}(i)$: neighborhood of node i
- σ: LeakyReLU activation

- Multi-head attention aggregates Kheads:

$$h_i' = \frac{1}{K}\sum_{k=1}^{K} h_i^{(k)}$$

2. MaterialsGCN

- Uses Graph Convolutional Networks (GCNConv).
- Node update rule:

$$h_i' = \sigma\left(\sum_{j \in \mathcal{N}(i)} \frac{1}{\sqrt{d_i d_j}} W h_j\right)$$

Where:

- d_i is the node degree
- W is a learnable weight matrix

3. MaterialsGraphSAGE

- Uses SAGEConv with mean aggregation:

$$h_i' = \sigma\left(W \cdot \mathrm{CONCAT}\left(h_i, \frac{1}{|\mathcal{N}(i)|}\sum_{j \in \mathcal{N}(i)} h_j\right)\right)$$

After message passing:

- Global graph embedding via mean pooling:

$$h_{\text{graph}} = \frac{1}{N}\sum_{i=1}^{N} h_i'$$

- Global features (u) are processed by an MLP:

$$u' = \mathrm{MLP}(u)$$

- Concatenated feature:

$$z = [h_{graph} \ || \ u']$$

- Final regression head predicts:

$$[\widehat{E}_{form}, \widehat{E}_{g}] = \text{MLP}(z)$$

b. Explanation Mechanism (Attention in GAT)
The attention mechanism in the GAT layers acts as an explainability tool. Each attention coefficient α_{ij} quantifies the importance of atom j to atom i. By visualizing or averaging α_{ij}, the model highlights which interatomic bonds most influence the predicted energy or band gap, providing a built-in interpretability similar to SHAP or LIME in tabular models.

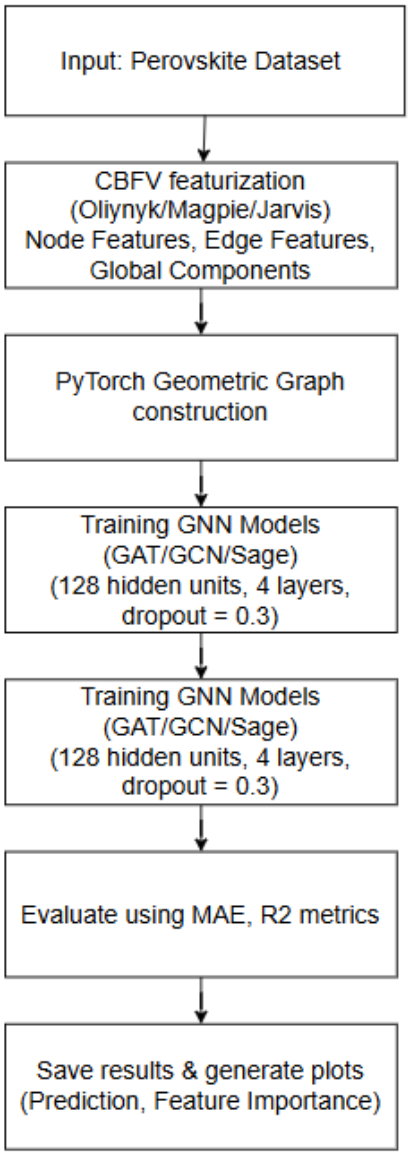
c. Training and Hyperparameter tuning
   i. Source of the dataset
   The project uses the Dataset "Perovskite.csv" for training the models. It contains material compositions, formation energy and band gap values for different materials.
   ii. Dataset description
      - Each entry is a crystalline compound along with its physiochemical properties and electronic properties.
      - The dataset was pre-processed and outliers were dropped. Other materials (i.e. materials not meeting the criteria in desirable properties) were also removed and a newer dataset called "Perovskite_data_cleaned.csv" is created.
      - CBFV Featurization is done over this using the presets – "Oliynyk, Magpie, Jarvis".
   iii. Configurations of the system used
      - Framework: PyTorch, PyTorch Geometric, Scikit-learn
      - System:
         - GPU (CUDA) optional
         - 8 GB RAM minimum
         - OS: Ubuntu / Windows

   iv. Computational Efficiency
   To reduce computational load:
      - Early stopping: stops training when validation loss no longer improves.
      - Learning rate scheduler: uses ReduceLROnPlateau for adaptive LR reduction
      - Gradient clipping: prevents exploding gradients.
      - Mini-batching: allows effective utilization of memory for graph datasets.

- Feature selection: employs only necessary CBFV descriptors for each predetermined.

d. Validation strategy
   - The dataset is split into training set (64%), validation set (16%) and testing set(20%)
   - The performance of the model will be evaluated using performance metrics such as MAE (Mean Absolute Error) and R2 (Coefficient of Restitution)
   - Separate metrics have been calculated for each of the prediction.

e. Flow-diagram



4. Results and Discussion
    a. Discussion of Model training, and validation plots (loss-convergence, how accurate the predictions are for the unseen dataset)

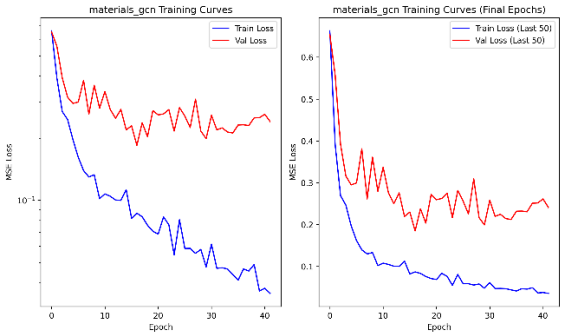


Fig 1: materials_gcn Training Curves

Fig 2: materials_gnn Training Curves

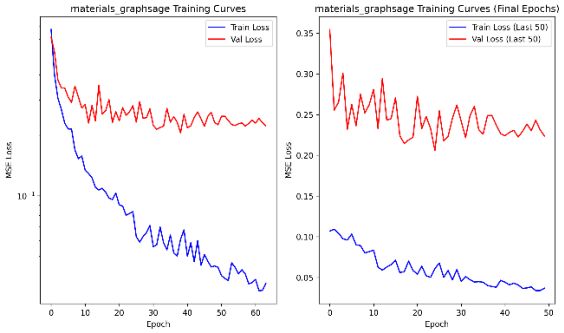


Fig 3: materials_graphsage Training Curves

The training and validation loss plots for the three models (GCN,GAT/GNN, GraphSage) show import insights into how well the models learns during training and how accurately the perform on unseen validation datasets.

- Loss Convergence Analysis
    - For all three models, the training loss (blue curve) decreases steadily over epochs, indication each model is successfully minimizing the error values.
    - The validation loss (red curve) plateaus after an initial drop for each model, meaning that after a point, further training is not reducing error values. This may be because of model reaching its optimal capacity.

GCN Predicted vs Actual Graphs:

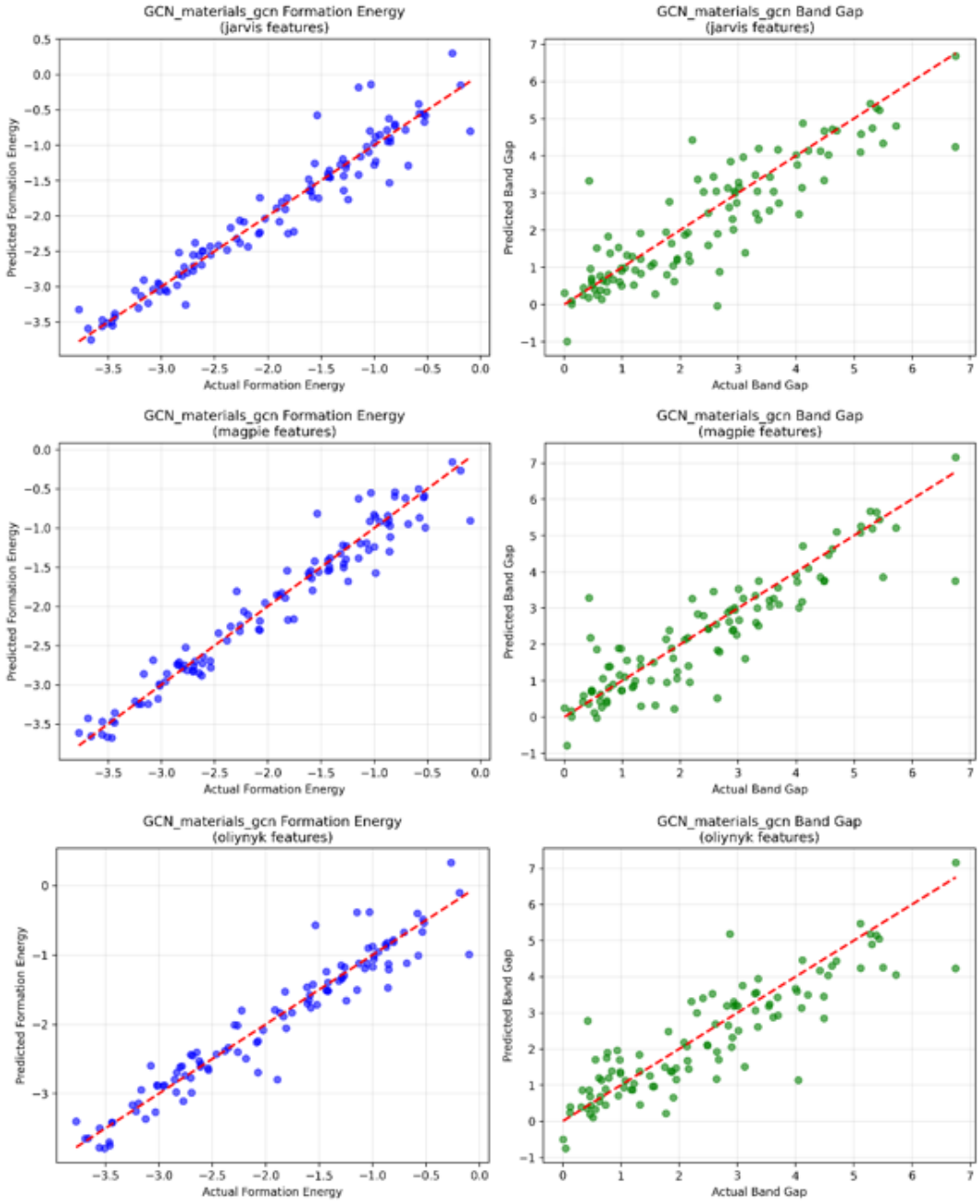


Fig 4: materials_gcn Predicted vs Actual Graph

Understanding: The predictions for Formation Energy (blue points) for the GCN model show a very strong correlation across all three feature sets. The data points for Jarvis and Magpie features are accurate, whereas in Oliynyk features, there is a slight bias as predicted values are falling just below the red line.

The predictions for Band Gap (green points) are comparable to others, with a moderate amount of scatter. The scatter is quite prominent especially in Oliynyk features. This suggests that GCN are less precise in predicting Band Gap.

GNN/GAT Predicted vs Actual Graphs:

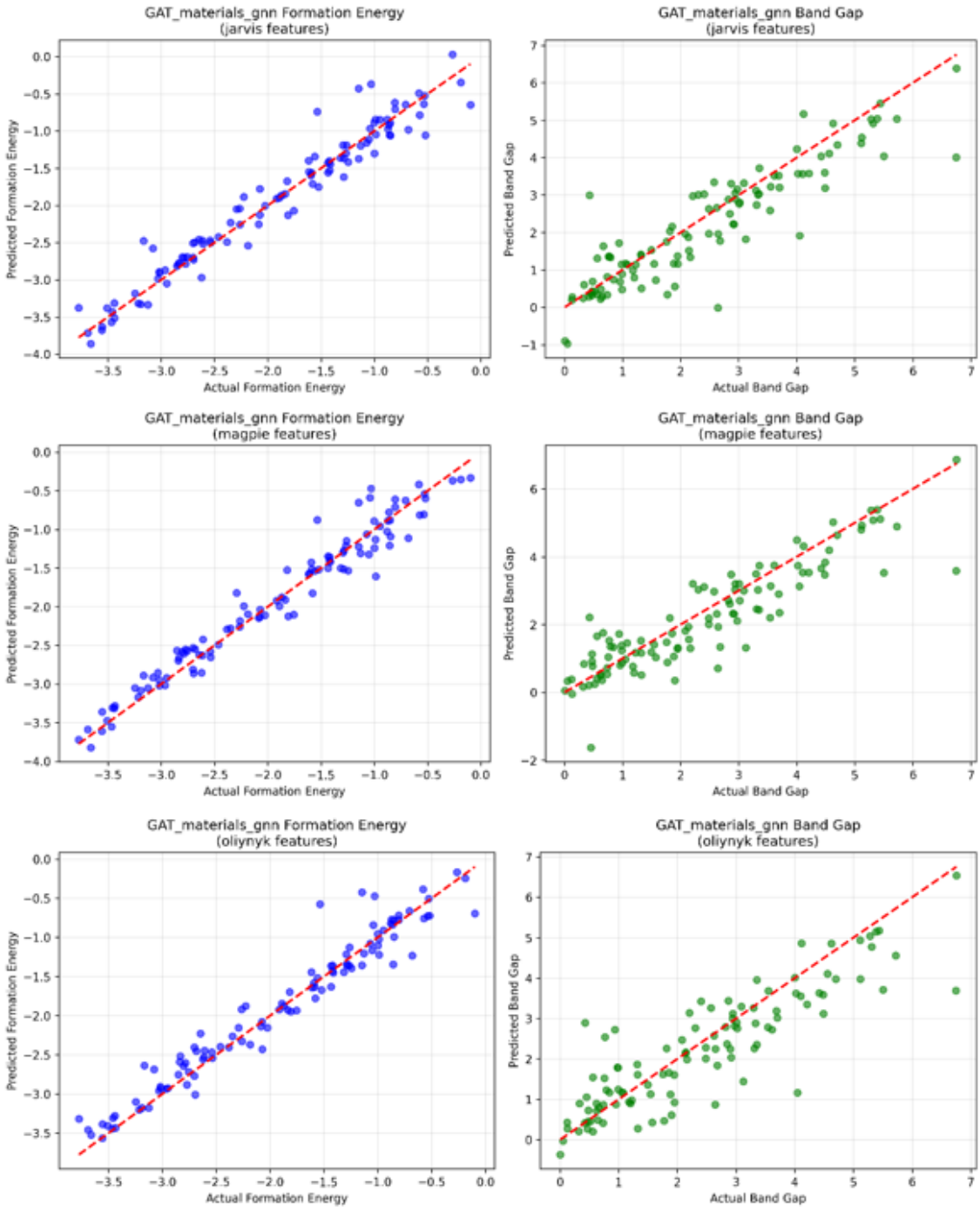


Fig 5: materials_gnn Predicted vs Actual Graph

Understanding: Similar to GCN, the GAT model shows excellent predictive accuracy for Formation energy, as seen by the tight clustering of data points around the red line.

The predictions for Band Gap (green points) show a good correlation, but there is noticeable scatter. The scatter seems particularly significant for low band gap (< 1eV) and high band gap (>5eV), suggesting a slight struggle to capture the variation at the extremes.

GraphSAGE Predicted vs Actual Graphs:

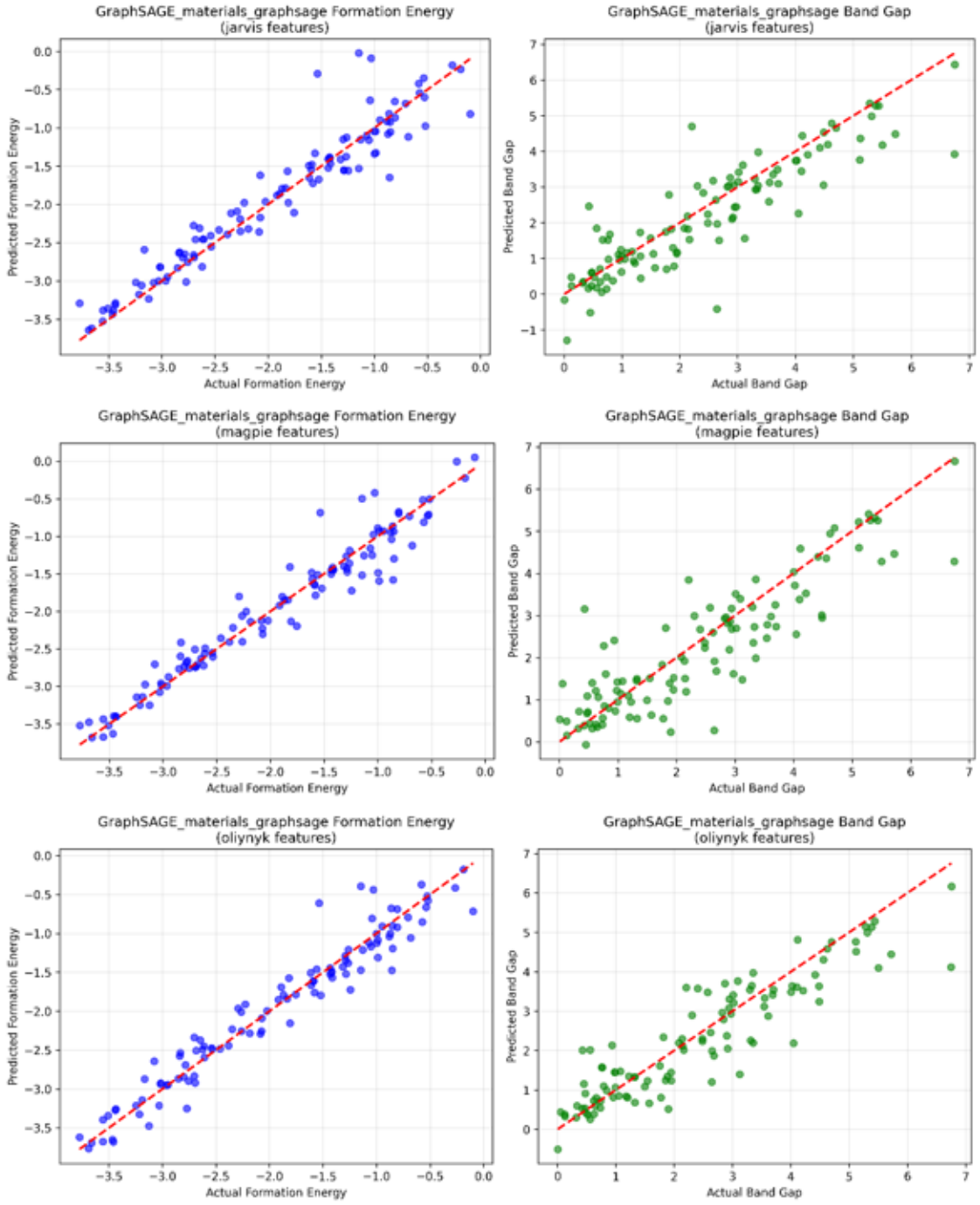


Fig 6: materials_graphsage Predicted vs Actual Graph

Understanding: The predictions for Formation Energy (blue points) show a linear correlation across all three feature sets. The data points are clustered around the y = x line, indicating the model is highly accurate in predicting Formation Energy.

The predictions for Band Gap (green points) also show a general strong positive correlation with actual values. This suggests that GraphSage are less precise in predicting Band Gap, particularly for higher Band Gaps (> 4eV), where more scatter is seen.

Comparative Summary:

In all three GNN models (GraphSAGE, GAT, GCN) and all three feature sets (Jarvis, Magpie, Oliynyk),

- Formation Energy is correctly and consistently predicted, and is highly correlated with actual values (tightly clustered about y = x). This means the structural and elemental information stored in the features works very well in predicting this thermodynamic property.
- Band Gap predictions have a good overall trend but are more scattered (lower precision) than Formation Energy. This indicates that Band Gap, a property of electronics, is harder to predict precisely with such combinations of GNN and features because it is very sensitive to subtle details in electronic structures.
- No single GNN and set of features are unequivocally superior for both properties. The performance for Formation Energy is great and similar for all variants. For Band Gap, the variations are slight, with all the models capturing similar amounts of scatter.

Prediction Metrics:

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

Fig 7: materials_gcn Prediction Metrics

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

Fig 8: materials_gnn Prediction Metrics

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

Fig 9: materials_graphsage Prediction Metrics

Across all three GNN architectures, the predictive models consistently demonstrate accurate performance in the case of Formation Energy (test R2 > 0.92) compared to Band Gap (test R2 = 0.75 - 0.82). The explicit featurizers (Jarvis, Magpie, Oliynyk) have offered little or no improvement for Formation Energy, and their impact on Band Gap prediction is mixed, where Magpie shows slight benefit in GCN model and Oliynyk in the case of GraphSAGE. The overall best results for Formation Energy are R2 = 0.947 and for Band Gap is R2 = 0.768.

5. Conclusions
   This work was able to explore the use of three different Graph Neural Network (GNN) architectures—GNN (Graph Attention Network, GAT in the plots), GraphSAGE, and GCN—to predict two important material properties: Formation Energy and Band Gap. The performance of these models was tested in a systematic manner using various external material feature sets (Jarvis, Magpie, and Oliynyk) along with the intrinsic features extracted by the GNNs.

Summary of Work

We trained and tested the GAT, GraphSAGE, and GCN models on a materials dataset, predicting Formation Energy (a thermodynamic property) and Band Gap (an electronic property). The performance was evaluated on the standard regression metrics (MAE and R2) and plotted by way of parity plots that graphically assessed actual versus predicted. This holistic evaluation enabled direct comparison of model architectures and feature engineering approaches.

Quantitative Findings:

- High Formation Energy Predictability: All the GNNs performed extremely well for Formation Energy with consistently high test R2 values of more than 0.92 and low MAE values (0.15 - 0.17). This shows that the GNNs, using structural and elemental descriptors, are very good at describing the intrinsic thermodynamic stability of materials.
- Lower Band Gap Predictability: Band Gap prediction was much tougher, with the test R2 values congregating lower between 0.74 and 0.82. The MAE values (0.50 - 0.60) were roughly three times larger than those for Formation Energy, corresponding to the higher level of complexity in describing this electronic property.
- Minimal Feature Impact: For the top-performing Formation Energy predictions, adding explicit featurizers (Jarvis, Magpie, Oliynyk) provided no consistent improvement over the GNNs' own representations learned. For Band Gap, the influence was small, with GraphSAGE + Oliynyk and GCN + Magpie demonstrating only modest edges.

Qualitative Findings:

Formation Energy (Visual Accuracy): The blue data points for Formation Energy were very tightly bunched along the perfect y = x red line for all models and features. This reflects high accuracy and low bias, graphically verifying the super R2 values.

Band Gap (Visual Scatter): The green data points for Band Gap had significant scatter away from the y = x line, especially in the extremes (low and high Band Gap values). This visual dissemination easily justifies the lower R2 values and the larger MAEs, showing that the models are able to pick up on the general trend but fail at high-precision prediction, particularly for extreme or outlier materials.

Model Robustness: All three GNN models (GAT, GraphSAGE, GCN) showed equivalent performance on both properties, indicating that the quality of the input data and the inherent complexity of the property (Band Gap vs. Formation Energy) were more predictive of the final accuracy than the particular message-passing mechanism of the GNN.