Bridging Feature-Based Models and Graph Neural Networks: A Hybrid Approach for Accurate and Interpretable Materials Modeling

Rogério Almeida Gouvêa[1,*], Pierre-Paul De Breuck[2,†], Tatiane Pretto[1], Gian-Marco Rignanese[2,*], Marcos José Leite dos Santos[1]

[1] Laboratory of Applied Materials and Interfaces, Federal University of Rio Grande do Sul, Porto Alegre, RS 91501-970, Brazil.

[2] Institute of Condensed Matter and Nanosciences, Université catholique de Louvain, Louvain-la-Neuve, Belgium.

*Corresponding authors: rogeriog.em@gmail.com; gian-marco.rignanese@uclouvain.be;

ABSTRACT

This study introduces an innovative method to enhance feature-based machine learning models in materials science by using graph neural networks (GNNs) as advanced featurizers. Our approach blends the interpretability of traditional models with GNNs' predictive power, creating a versatile tool for materials informatics. Two methods are developed to incorporate GNNs as featurizers: using GNNs to obtain latent representations of electronic structure descriptors and integrating pretrained GNNs for different target properties. When combined with standard chemical features, these GNN-enhanced models improve prediction accuracy by up to 44.2%. Evaluated on perovskite heat of formation, this method reduces the mean absolute error significantly, surpassing both standard performance with the feature-based model MODNet or with the GNN model MEGNet with default hyperparameters. It also shows promising results for complex tasks, such as predicting convex hull distances and band gaps. SHAP analysis and surrogate models are used to extract chemical insights, reinforcing the model's interpretability. This approach advances materials modeling, enabling more accurate, efficient, and interpretable predictions for targeted materials design.

Keywords: Feature-based machine learning, MODNet, graph neural networks, materials informatics, interpretability.

† Current address: Ruhr-Universität Bochum, Universitätsstr. 150, 44801 Bochum, Germany.

Introduction

Machine learning has revolutionized materials science, accelerating material discovery and property optimization across various domains[1–3]. The two prominent approaches in this field are the feature-based and graph neural network (GNN) models, each with distinct advantages and limitations[4,5]. Feature-based models rely on predefined descriptors like elemental properties, geometric features, and electronic structure information. They are highly interpretable and effective with small datasets, offering insights into structure-property relationships[6,7]. These models adapt well to custom tasks in experimental settings such as nanocrystal research[8], catalysis[9], and organic photovoltaics[10]. In contrast, GNN models represent materials as graphs, capturing structural information through message passing and learning deep representations without predefined descriptors. This often results in more accurate predictions for complex materials but demands greater computational resources and data for training[11,12]. GNNs are particularly effective in large-scale screening of materials due to their efficient computation and local information aggregation[13].

Feature-based models with extensive suites like MatMiner[7] face time-consuming featurization and the curse of dimensionality, particularly with complex descriptors like the Orbital Field Matrix (OFM)[14] and the Smooth Overlap of Atomic Positions (SOAP)[15]. GNN models, while computationally efficient, often lack interpretability and struggle with smaller datasets. In this work, we address these challenges by proposing a hybrid approach that combines traditional, chemically intuitive descriptors with GNN-derived features obtained through proxy Graph Neural Network (pGNN) featurizers. The latter integrate pretrained GNN models to extract insightful chemical information, enhancing the interpretability and efficiency of feature-based models on larger datasets.

GNN-derived features are obtained in two ways. First, electronic structure descriptors or other chemical features, such as those from MatMiner, are encoded into a latent space using an

autoencoder. A GNN trained to reproduce these latent representations directly from the input structures is then used. Alternatively, we leverage pretrained GNNs tailored to specific target properties. The interpretability of models based on pGNN featurizers is retained through techniques such as Shapley Additive Explanations (SHAP), which correlates GNN-derived features with chemically intuitive descriptors. These GNN-derived features are incorporated with minimal computational overhead using pretrained models. Comprehensive implementation details are available in the Methods section and in the *Supporting Information*.

Building upon this foundation, our hybrid featurizer, named OMEGA, integrates features from MatMiner with three distinct GNN-derived featurizers derived from the pGNN framework to enhance material property predictions. These featurizers rely on different GNN models as illustrated in Fig. 1: (1) a latent-OFM GNN, (2) pretrained MEGNet models from the Materials Virtual Lab (MVL), and (3) an Adjacent, task-specific GNN model. The latent OFM GNN captures electronic structure interactions by directly deriving encoded OFM features from the material structures. The pretrained MEGNet models from MVL contribute additional chemical insights, including features related to the formation energy and elastic constants. The adjacent GNN model is specifically trained on the target dataset to capture task-specific data nuances, thereby improving prediction accuracy. Once trained, these GNN featurizers can be readily applied to other tasks without requiring additional training, leveraging their speed and adaptability.

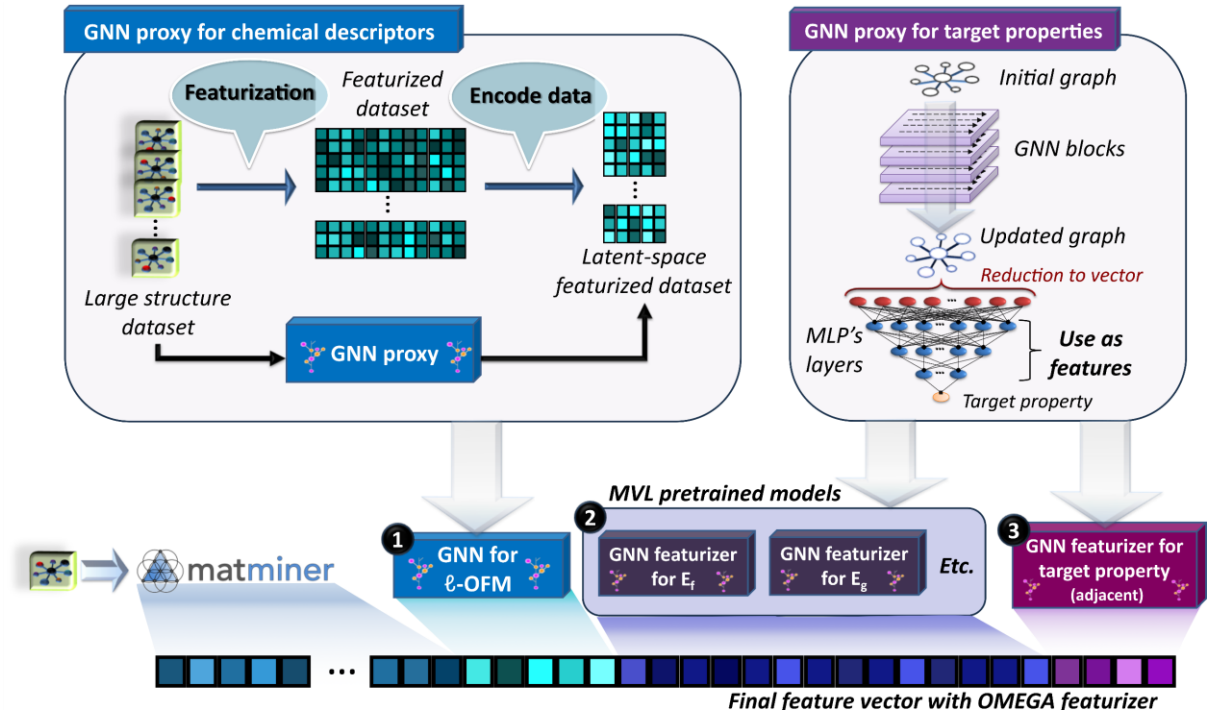

Fig. 1 | Overview of the methodology for leveraging GNN models to produce the OMEGA featurizer. The left side of the diagram illustrates how a GNN model acts as a proxy to generate the latent-space representation of a chemical descriptor, specifically the OFM features (1). The right side shows that, instead of using the predicted value directly from pretrained GNN models, values are extracted from the final multi-layer perceptron (MLP) regression layers. These values are obtained from either pretrained GNN models developed by the Materials Virtual Lab (2) or a GNN model specifically trained for the intended task (3).

Using the OMEGA featurizer with the feature-based MODNet model (MODNet@OMEGA) reduces the mean absolute error (MAE) for predicting the perovskite heat of formation by 44.2% compared to the baseline MODNet featurizer. Additionally, it is effective in more challenging tasks, such as predicting band gaps and stability for halogenated compounds in the Open Quantum Materials Database (OQMD).

Our hybrid approach bridges feature-based and graph-based methods, leveraging their strengths to develop more versatile and task-agnostic machine learning models in materials science. By enhancing accuracy, efficiency, and interpretability in property prediction, this model facilitates the integration of both experimental and simulated data. Moreover, it aligns with the growing demand for explainable AI[16,17], essential for the advancement of self-driving laboratories in materials discovery and optimization[18].

Results and discussion

The dataset of formation energy for perovskites named matbench_perovskites (MatBench v0.1), which measures the performance of a model in predicting the formation energy of various perovskites, is used as a proof-of-concept for testing all the new implementions related to the new featurizers. This task is selected because perovskites are a technologically important class of materials with complex compositions, making them a great candidate for demonstrating the effectiveness of our hybrid approach. Additionally, it represents the smallest benchmark task, with a dataset comprising 18,928 samples, where the performance gap between graph-based and feature-based models becomes significant. The benchmarked mean absolute errors (MAE) for MEGNet, MODNet, AutoMatMiner and RF/SCM-MagPie on this task are shown in Table 1, along with our own results for the MODNet and MEGNet model in this task.

Table 1 | Mean absolute errors (MAEs) for MatBench task of heat of formation of perovskites (matbench_perovskites) with different algorithms.

General Purpose Algorithm	MAE on task matbench_perovskites (eV)
MEGNet*	$0.0352\ (\pm 0.0016)$
MEGNet (this work, transferred embedding**)	$0.0685\ (\pm 0.0036)$
MEGNet (this work, no embedding)	$0.0840\ (\pm 0.0058)$
MODNet*	$0.0908\ (\pm 0.0028)$
MODNet (this work)	$0.0888\ (\pm 0.0025)$
AutoMatMiner*	$0.2005\ (\pm 0.0085)$
RF-SCM/MagPie*	$0.2355(\pm 0.0034)$

*Data were retrieved from *MatBench*[12] in May 2024.
**Elemental embedding transferred from the formation energy task in the MEGNet repository.

Using the default MatMiner featurizer in MODNet (all details are available in the *Supporting Information*), the MAE is 0.0888 eV for the matbench_perovskites. This result falls within the error margin of 0.0908 ± 0.0028 eV retrieved from the MatBench for MODNet. This value serves as the reference for all subsequent implementations to evaluate the

performance gain on this task. The MEGNet benchmarked MAE is substantially lower than what we achieved using the default configuration of the model, both with and without transferring elemental embedding from the model trained on larger datasets provided by the authors. This discrepancy arises due to several factors, including differences in hyperparameters such as the size of the MEGNet blocks, the dimensionality of the embedding vectors, and the inclusion of additional features like charge information[19]. Moreover, the benchmarked MEGNet model may have benefited from longer training times. However, we chose to stick with the default configuration for faster training when employed in our setup. Details of the hyperparameters used for MODNet and MEGNet models in this work are provided in the *Supporting Information*.

To systematically enhance our feature set for predicting the heat of formation of perovskites, we incrementally incorporate pGNN features, as illustrated in Fig. 2. The details about the methodology used for developing the GNN featurizers for OFM and MatMiner chemical descriptors (ℓ-OFM pGNN and ℓ-MM pGNN), including latent-space optimization, are provided in the *Supporting Information*. We first assess the performance of ℓ-OFM pGNN and ℓ-MM pGNN features compared to the baseline MatMiner featurizer with original OFM features. Next, we integrated pretrained GNN models from MVL, evaluating different multilayer perceptron layers and analyzing the effect of dataset size. Finally, we incorporated an adjacent pGNN featurizer (adj. pGNN), trained specifically on the matbench_perovskites task. Combining these featurizers as shown in Fig. 2 results in the hybrid featurizer, OMEGA. Additionally, we developed OMEGAfast, a variant that utilizes ℓ-MM pGNN features instead of the original MatMiner features for faster featurization.

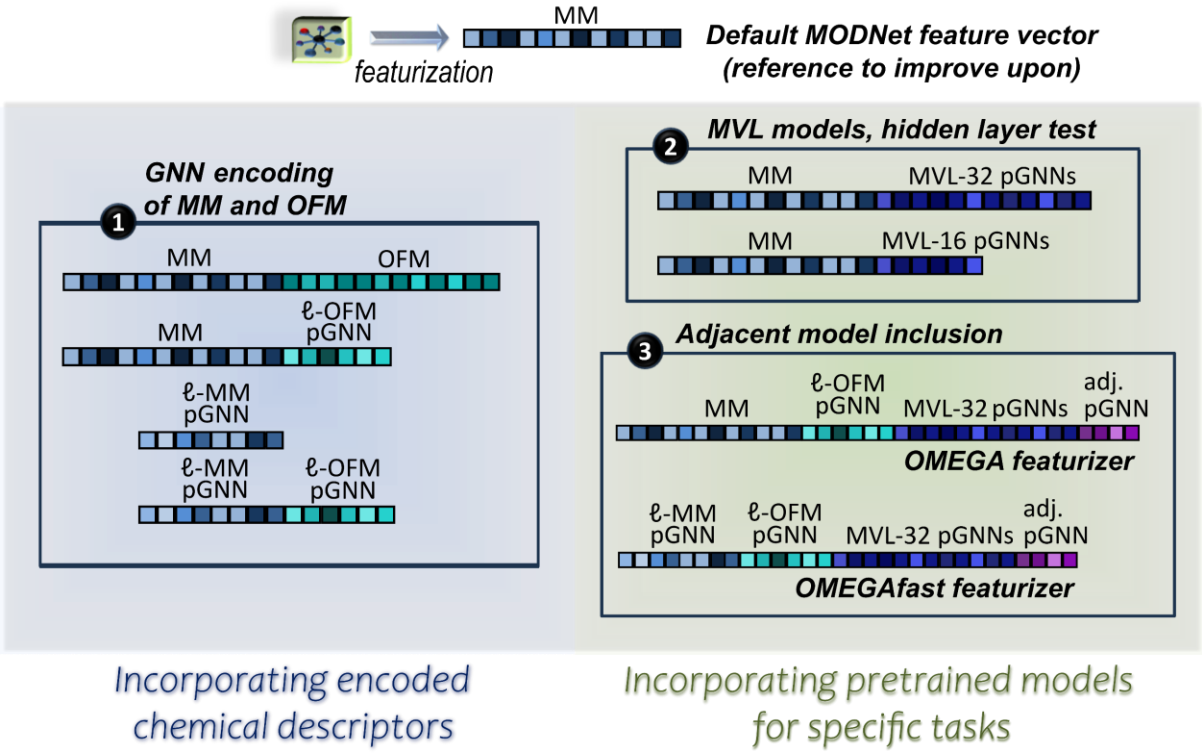

Fig. 2 | Schematic of the OMEGA and OMEGAfast featurizer development for improved prediction of perovskite heat of formation. The steps include: (1) comparing OFM and MM descriptors with GNN-derived latent features; (2) testing multilayer perceptron layers from pretrained MVL pGNN models; and (3) finalizing with an adjacent pGNN featurizer trained specifically on the heat of formation task.

GNN featurizers for latent-space representations of chemical descriptors

We first benchmark the addition of OFM features to the original MatMiner features, as presented in Table 2. This leads to a performance boost compared to using solely the default MatMiner features for prediction. Over 900 new features including more specific chemical information about the orbital interactions in each structure which proves to be beneficial to the model accuracy. It is noteworthy that the curse of dimensionality is mitigated through MODNet's robust feature selection algorithm and subsequent hyperparameter optimization through a genetic algorithm to choose an optimal subset of features for the given problem.

In Table 2, the models including GNN-derived latent MatMiner features, ℓ-MM pGNN, show that increase MAE by 0.025 eV most probably due to reconstruction errors. However, these models still outperform Automatminer and random forest benchmarks (Table 1) and

allow faster featurization. For the OFM features, the GNN-derived latent representation performs nearly as well as the original, with only a 0.0051 eV decline. Combining latent features from both pGNN models slightly reduces the MAE, highlighting the potential benefits of integrating multiple latent representations of chemical descriptors.

These results emphasize the effectiveness of our proposed proxy GNN featurizers in capturing essential chemical information, even in the presence of reconstruction challenges, while also significantly reducing the computation time and making the feature-based models more efficient for large-scale applications. By further refining the models, such as training on larger, carefully curated datasets, we may mitigate reconstruction errors and enhance the pGNN featurizers' ability to identify chemical patterns.

Table 2 | Mean absolute errors (MAEs) for MODNet models on matbench_perovskites task comparing the inclusion of latent features originally obtained from the autoencoder and through the GNN featurizers. The relative MAE deviation from the default MatMiner featurizer in MODNet is reported in parentheses.

Features	MAE (eV)
Default MatMiner (MM)	0.0888
MM + original OFM	0.0751 (−15.3%)
ℓ-MM	0.0793 (−10.7%)
ℓ-MM pGNN	0.1052 (+18.5%)
MM + ℓ-OFM	0.0743 (−16.2%)
MM + ℓ-OFM pGNN	0.0794 (−10.6%)
ℓ-MM pGNN + ℓ-OFM pGNN	0.0973 (+9.6%)

MVL pGNN featurizers from pretrained models

To incorporate pretrained GNN models from MVL as features, we extract the values from the last layers of the MLP regression head of the MEGNet model architecture. Table 3 presents

a performance comparison for the matbench_perovskites task, incorporating the hidden layers with 32 neurons (referred to as MVL-32 pGNNs) and the layers with 16 neurons (referred to as MVL-16 pGNNs). Additionally, we conduct assessments on randomly selected subsets comprising 5000 samples and 1000 samples from the initial matbench_perovskites dataset to verify the consistency of our findings across smaller datasets.

Our analysis reveals a consistent enhancement in performance with the inclusion of the MVL-32 pGNNs featurizer over the MVL-16 pGNNs featurizer, irrespective of the dataset size. This improvement is attributed to a more general latent-space representation in the earlier layers of the model, which MODNet can effectively leverage. Notably, the percentage reduction in MAE compared to the exclusive use of MatMiner features increases as the dataset size decreases. This underscores the transfer learning essence of this technique, transferring pre-acquired chemical knowledge from larger datasets to enhance performance on small datasets.

Table 3 | Mean absolute errors (MAEs) for MODNet models on the matbench_perovskites task and subsets comparing the inclusion of features from pre-trained MEGNet models distributed by Materials Virtual Lab. *N* represents the size of the dataset used for the prediction. The relative MAE deviation from the default MatMiner featurizer in MODNet is reported in parentheses for each task.

Features	Task		
	matbench perovskites (N=18,928)	matbench perovskites (N=5,000)	matbench perovskites (N=1,000)
	MAE (eV)	MAE (eV)	MAE (eV)
Default MatMiner (MM)	0.0888	0.1667	0.2802
MM + MVL-16 pGNNs	0.0752 (−15.3%)	0.1202 (−27.9%)	0.1862 (−33.5%)
MM + MVL-32 pGNNs	0.0726 (−18.2%)	0.1167 (−30.0%)	0.1749 (−37.6%)

We proceed to examine the effects of evaluating ℓ-MM and ℓ-OFM featurizers with the best-performing MVL-32 pGNNs featurizer, as outlined in Table 4. It can be seen that, despite

the substantial contribution of the MVL pGNNs to the accuracy, the addition of ℓ-OFM features brings further improvement, highlighting a synergistic effect when these set of descriptors are integrated. The combined latent representations result in a total percent reduction of 26.5% in MAE over the default MatMiner featurizer in MODNet. In addition to enhancing performance, pGNN featurizers integrate multiple models efficiently and provide reusable latent-space representations for diverse tasks.

Table 4 | Mean absolute errors (MAEs) for MODNet models on matbench_perovskites task comparing the inclusion of OFM latent features and both OFM latent features and MVL-32 GNNs features. The relative MAE deviation from the default MatMiner featurizer in MODNet is reported in parentheses.

Features	MAE (eV)
Default MatMiner (MM)	0.0888
MM + ℓ-OFM	0.0743 (−16.2%)
MM + MVL-32 pGNNs	0.0726 (−18.2%)
MM + ℓ-OFM + MVL-32 pGNNs	0.0629 (−29.1%)
ℓ-MM	0.0793 (−10.7%)
ℓ-MM + ℓ-OFM	0.0728 (−18.0%)
ℓ-MM + MVL-32 pGNNs	0.0729 (−18.0%)
ℓ-MM + ℓ-OFM + MVL-32 pGNNs	0.0653 (−26.5%)

In contrast to features based on chemical principles, the latent-space features from pGNN lack immediate interpretability. **Fig. 3** showcases some of the most relevant features in the MODNet model with the set of features "MM + ℓ-OFM + MVL-32 pGNNs" determined through SHAP plot, and their relationship to interpretable features with the highest SHAP value, determined through surrogate models. The methodology to recover interpretability from GNN features is provided in Methods. Additionally, extended SHAP summary plots for each tested model in this work are presented in the *Supporting Information*. The MEGNet pretrained

features of the formation energy model are the most relevant for the model prediction. When decomposed into chemical descriptors, we observe that more electronegative elements (*MagpieData minimum Electronegativity*, in **Fig. 3**) tend to increase the heat of formation, just as increased difference in the number of electrons in the valence shell (*MagpieData avg dev NUnfilled*) or elemental ground-state band gap (*MagpieData maximum GSbandgap*). As expected, the heat of formation of perovskites also increases when the interaction of orbitals s^2 and p^4 is present (*OFM:* s^2-p^4), characteristic of many oxide perovskites, conversely to when two pnictogen elements are present (*OFM:* p^3-p^3), which creates weaker chemical bonds. The observation on combining pnictogens had already been made for perovskites in the high-throughput screening performed by Schmidt et al. (2021), in which, no system alloying two pnictogens presented a decomposition energy lower than 100 meV/atom. The same effect is seen when Voronoi distances increase (*mean Voro dist minimum*). This occurs when larger A-site cations are present, which usually indicates a higher stability in perovskites[21]. Among the MatMiner features ranking higher in the model output, we observe that the presence of *d* orbitals tends to reduce the heat of formation, just as the presence of transition metals in general. A higher melting temperature, characteristic of higher binding energy, also correlates with a higher heat of formation. Analyzing the decomposition of the ℓ-OFM features, a couple observations can be drawn, such as mixed anion perovskites inducing weaker bonds (*OFM:* p^4-p^3) just as in the case of halide perovskites in general (*OFM:* p^5-s^2).

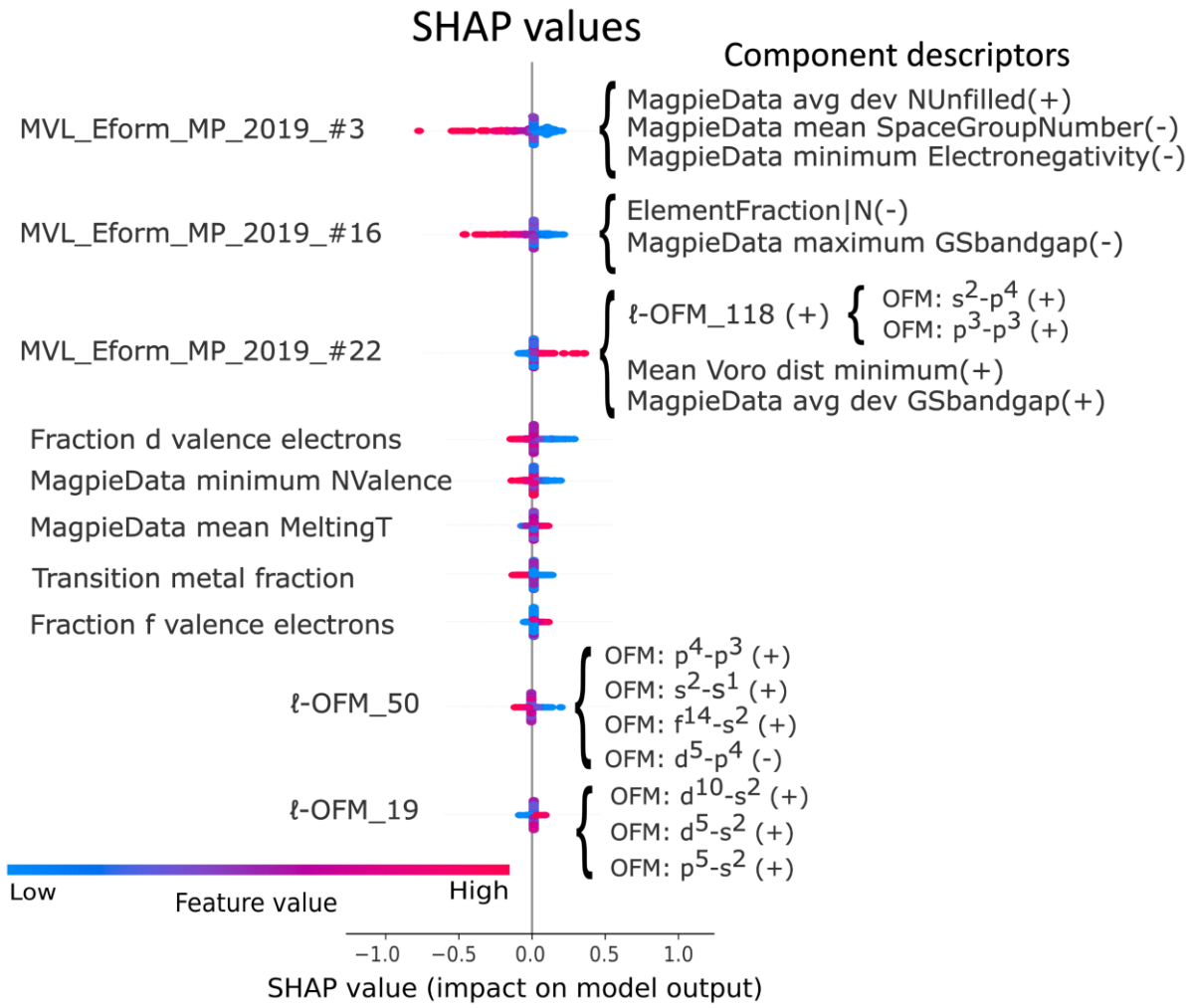

Fig. 3 | SHAP analysis of top features in MODNet model for perovskite formation energy with "MM + ℓ-OFM + MVL-32 pGNNs" features. MVL-32 pGNNs and ℓ-OFM features are decomposed into original MatMiner and OFM features, (+) indicates proportional variation and (-) indicates inversely proportional variation to the encoded features.

Adjacent pGNN featurizer and final featurizers

The latest addition to the proposed OMEGA featurizer is an adjacent pGNN featurizer, derived from a model trained on the target property, specifically the heat of formation of perovskites in our test case (see Methods for details). While this requires training an additional model, it leverages the flexibility of a GNN to enhance accuracy when sufficient training data is available. In Table 5, we integrate the adjacent pGNN features with the other pGNN featurizers for predicting the perovskite heat of formation, evaluating both scenarios: retaining the MatMiner features or using their latent representation via the ℓ-MM pGNN featurizer. The results show that each added pGNN featurizer improves the model's accuracy. The adjacent

model decreases the MAE by 0.0188 eV compared to adding only ℓ-OFM pGNN and MVL-32 pGNNs to the original MatMiner features. Additionally, using GNN-generated latent-space MatMiner features (ℓ-MM pGNN) reduces the MAE by 0.0227 eV, achieving nearly the same accuracy as when using MatMiner featurization. This indicates that the adjacent pGNN features dominate the predictions. However, a synergetic effect is clear when we compare the result to the MAE of the GNN model alone as presented in Table 1, which achieved a 0.0685 eV MAE, 0.019 eV (27.7 %) higher average error than our MODNet@OMEGA model.

In Fig. **4**, SHAP value analysis of the most important features for MODNet@OMEGA model shows that MatMiner, ℓ-OFM, and MVL-32 pGNNs features contribute similarly to the MODNet@"MM + ℓ-OFM + MVL-32 pGNNs" model. However, adjacent pGNN features dominate the model, correlating with more subtle patterns like geometrical fingerprints and Coulomb matrix eigenvalues, attributed to the targeted GNN model's ability to exploit highly non-linear relationships for enhanced accuracy.

Our results demonstrate that the proposed design allows for the exploration of intricate patterns for enhanced GNN accuracy through combined training with interpretable chemical descriptors and SHAP value analysis. This approach provides greater interpretability while approaching benchmarked GNN accuracy. Furthermore, using GNN models as proxies to generate latent-space features for MatMiner and OFM descriptors significantly reduces the computational cost of featurization, enhancing viability for high-throughput materials screening.

Table 5 | Mean absolute errors (MAEs) for MODNet models on matbench_perovskites task comparing the inclusion of all GNN featurizers over original MatMiner features and over GNN generated ℓ-MM features. The relative MAE deviation from the default MatMiner featurizer in MODNet is reported in parentheses.

Features	MAE (eV)
Default MatMiner (MM)	0.0888
MM + ℓ-OFM pGNN	0.0794 (−10.6%)
MM + ℓ-OFM pGNN + MVL-32 pGNNs	0.0683 (−23.1%)
OMEGA (MM + ℓ-OFM pGNN + MVL-32 pGNNs + adj. pGNN)	0.0495 (−44.2%)
ℓ-MM pGNN	0.1052 (+18.5%)
ℓ-MM pGNN + ℓ-OFM pGNN	0.0973 (+9.6%)
ℓ-MM pGNN + ℓ-OFM pGNN + MVL-32 pGNNs	0.0726 (−18.2%)
OMEGAfast (ℓ-MM pGNN + ℓ-OFM pGNN + MVL-32 pGNNs + adj. pGNN)	0.0499 (−43.8%)

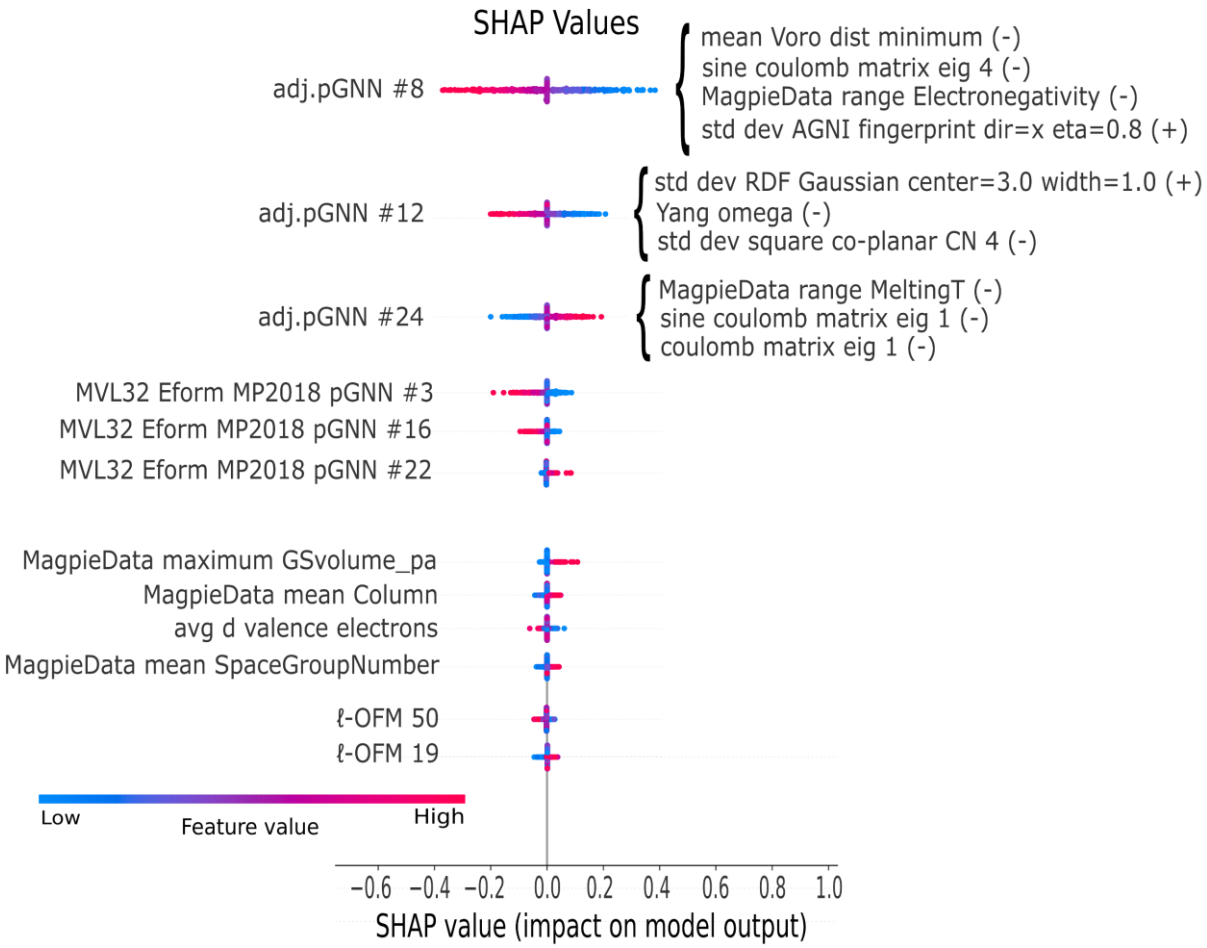

Fig. 4 | SHAP analysis of selected top features in MODNet model for perovskite heat of formation with OMEGA features. Adjacent GNN model features are decomposed into original MatMiner and ℓ-OFM pGNN features, where a few are shown, (+) indicates proportional variation and (-) indicates inversely proportional variation to the encoded features.

Validation of the hybrid model in advanced tasks

To further evaluate the proposed method, we conduct two additional tasks: predicting the decomposition energy (stability) and band gaps of halogen-containing materials from the OQMD dataset. We filter the OQMD dataset to include only structures containing halogens (F, Cl, Br, I). After preprocessing, structures with stability above 2.9 eV/atom (0.1% of the dataset) are removed, resulting in 31,271 samples for the stability prediction task. This dataset is further refined to 8,518 structures with band gaps above 0.5 eV for the band gap prediction task.

These tasks are more challenging than our proof-of-concept task on perovskites. Since the GNN featurizer for OFM and the pre-trained models from MVL have all been trained on the Materials Project dataset, which has limited information on halides[22], we anticipate lower

generalization for these datasets. In particular, estimating the decomposition energy is notoriously difficult and presents a current challenge on materials screening[23]. The results are presented in Table 6, where we observe minimal influence on stability predictions when including the ℓ-OFM pGNN featurizer. However, significant improvement is seen when including MVL-32 pGNNs, further enhanced by the addition of the adjacent pGNN features. For the band gap task, ℓ-OFM pGNN still fails to provide significant improvement to the model, but MVL-32 pGNNs and adjacent pGNN featurizers compensate with a notable enhancement.

Table 6 | Mean absolute errors (MAEs) for MODNet models on the prediction tasks of the convex hull distance and the band gap for the subset of halogen-containing materials from OQMD, comparing the inclusion of all pGNN featurizers over original MatMiner features. The relative MAE deviation from the default MatMiner featurizer in MODNet is reported in parentheses for the different tasks.

Features	Task	
	OQMD halogen E_{hull} (N=31,271)	OQMD halogen E_g (N=8,518)
	MAE (eV)	MAE (eV)
Default MatMiner (MM)	0.0556	0.4557
MM + ℓ-OFM pGNN	0.0561 (+0.8%)	0.4501 (−1.2%)
MM + ℓ-OFM pGNN + MVL-32 pGNNs	0.0538 (−3.2%)	0.3835 (−15.8%)
OMEGA	0.0519 (−6.7%)	0.3784 (−17.0%)

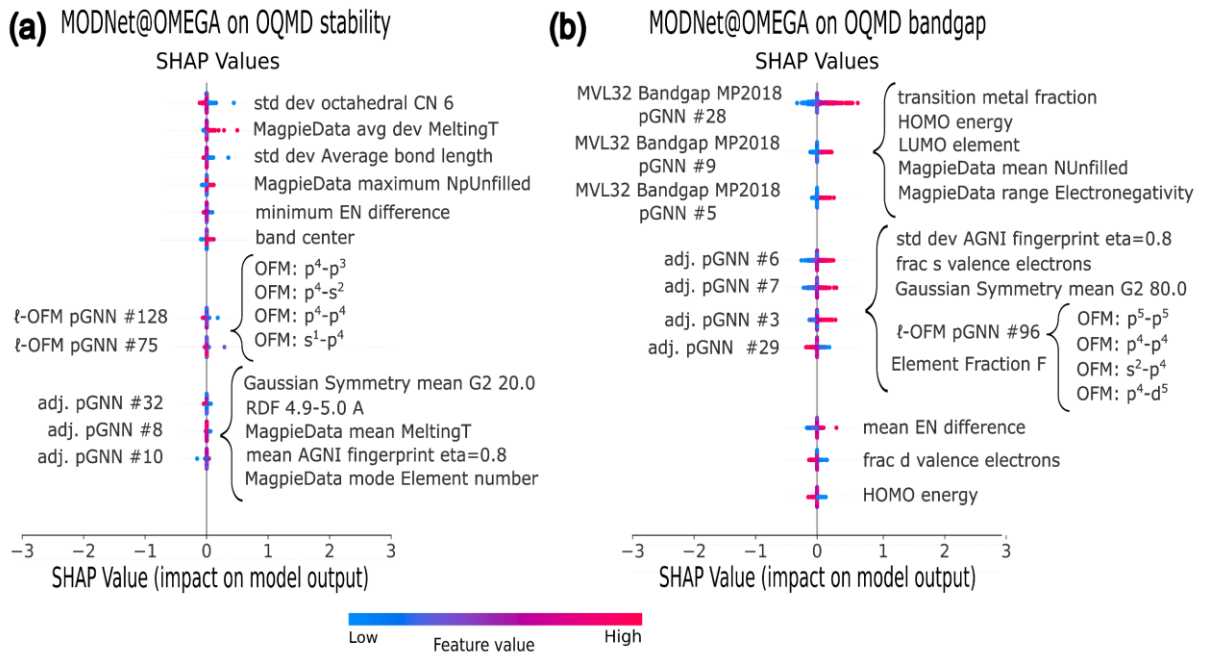

Fig. 5 | SHAP analysis of selected top features in MODNet model for OQMD halogen task for stability (a) and band gap (b) with OMEGA features. Adjacent model and MEGNet ℓ-OFM features are decomposed into chemical descriptors, which a few with highest impact on the group of features are shown.

In

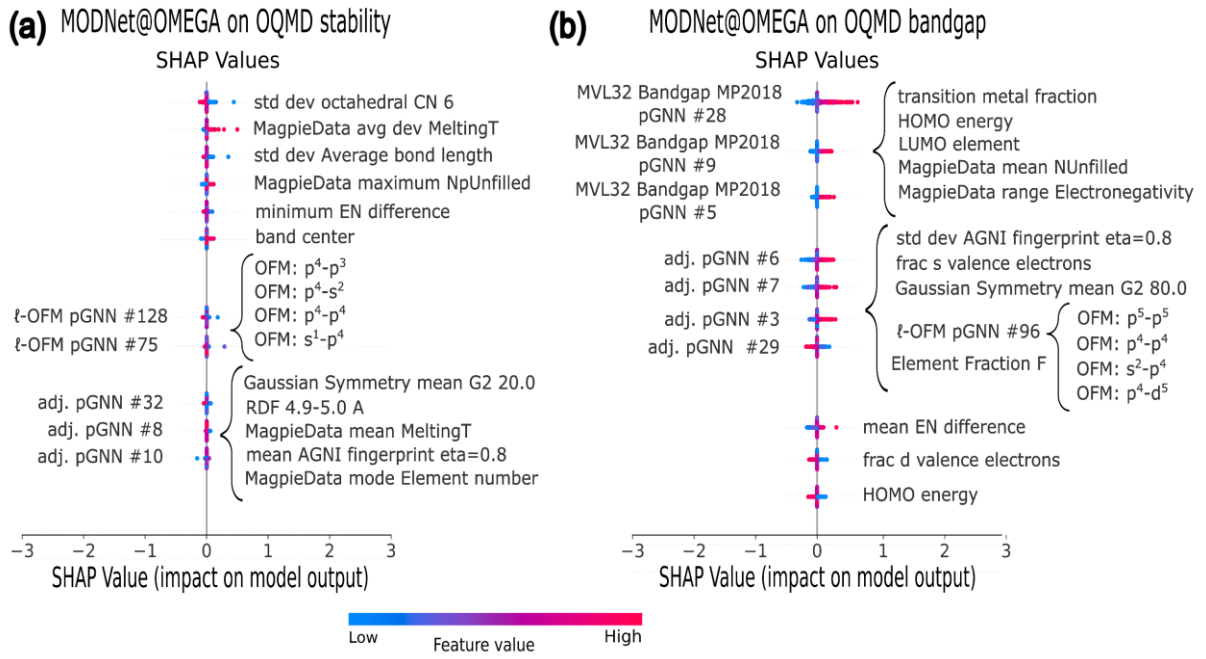

Fig. 5, the most important features for the MODNet@OMEGA model are displayed. For stability prediction (

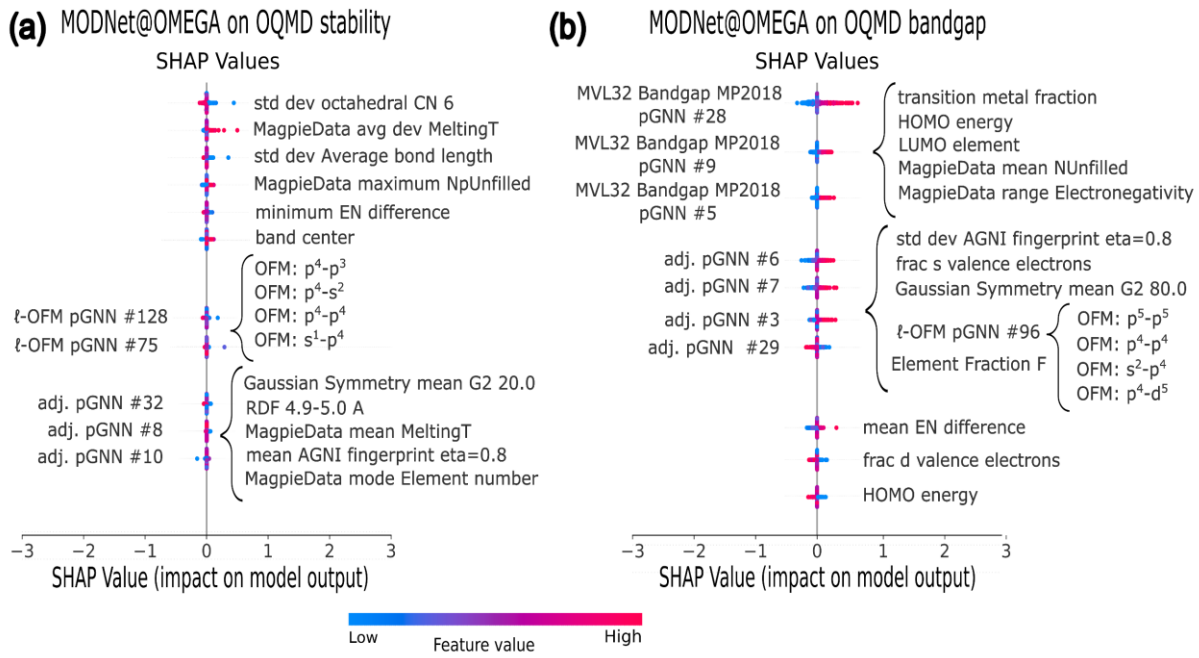

Fig. 5a), the original MatMiner features have a stronger influence than the additional OMEGA features, reflecting only a 6.7% MAE reduction compared to the perovskite task. This highlights the challenge of predicting stability from material structures, even with advanced GNN models[24]. The SHAP analysis shows an homogeneous feature importance with no dominant features, as further detailed in **Error! Reference source not found.**. This indicates a need for developing new material descriptors that better correlate with stability. Significant importance is assigned to geometric descriptors (e.g., geometrical fingerprints, symmetry, bond lengths) and traditional chemical descriptors (e.g., electronegativity difference, valence orbital filling, estimated melting temperature), consistent across MatMiner and adjacent model decompositions. Additionally, MatMiner's band structure featurizer, including elemental electronegativity, and specific ℓ-OFM features related to chalcogenides are prominent, supported by the prevalence of chalcogenides in inorganic materials.

For the band gap prediction task (

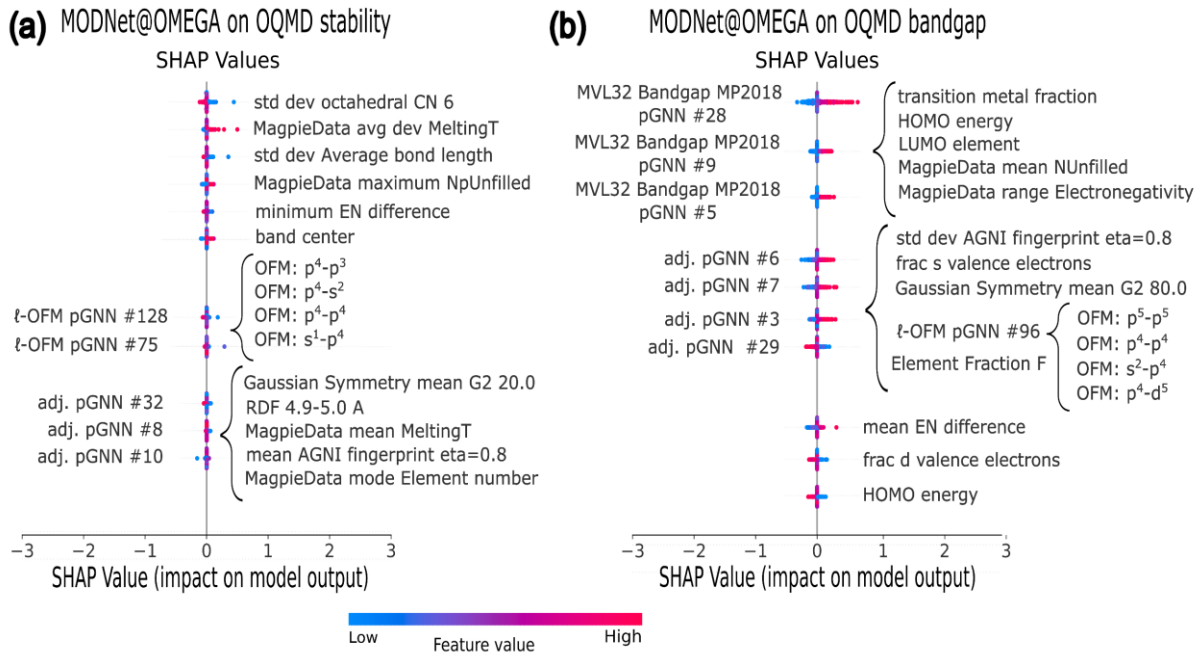

Fig. 5b), pGNN features are predominant. MVL pGNN features related to band gap regression are highly ranked, correlating with chemical descriptors such as electronegativity difference, HOMO/LUMO energies, valence band filling, and transition metal d-orbital filling. The adjacent model features also reveal similar chemical descriptors and additional geometric descriptors, enhancing the model accuracy as shown in Table 6 (see **Error! Reference source not found.** for extended SHAP plots).

In summary, our results validate the integration of GNN models as proxy featurizers in feature-based models, enhancing their performance on larger datasets. The OMEGA featurizer, utilizing featurizers from the pGNN tool reduce the MAE for perovskite heat of formation by 44.2% compared to the default MODNet featurizer, achieving an accuracy close to benchmarked GNN models. Additionally, the hybrid featurizer demonstrated generalizability in predicting convex hull distance and band gaps for halogen-containing materials in the OQMD dataset, particularly excelling in band gap predictions. Future studies could incorporate pGNN featurizers for sophisticated descriptors like SOAP and employ advanced GNN models such as ALIGNN[25] to further enhance predictive capabilities.

Moreover, this approach bridges the interpretability gap between feature-based models and highly accurate yet less interpretable GNNs. By employing SHAP analysis and surrogate models to evaluate the feature importance, we extract meaningful chemical insights from the GNN features. This enables the screening of vast chemical spaces and supports chemically guided active learning through inherent interpretability. In conclusion, integrating GNN features into feature-based models provides a versatile and powerful method to enhance predictions by leveraging pretrained GNN knowledge, reducing featurization costs, and improving model accuracy while retaining partial interpretability through techniques like SHAP analysis. This approach facilitates more accurate, efficient, and interpretable materials discovery using feature-based machine learning models.

Methods

GNN featurizers implementation

Our hybrid approach leverages the **pGNN** (Proxy Graph Neural Network) featurizer tool, implemented in Python, to enhance material property predictions by integrating pretrained Graph Neural Networks (GNNs) with traditional chemically intuitive descriptors. The pGNN package, available at https://github.com/rogeriog/pGNN, offers significant flexibility and modularity, allowing extraction of features from different layers of pretrained models and the incorporation of other GNN models as needed. This section outlines the detailed implementation and evaluation procedures for each of the pGNN featurizers utilized, along with their integration into our predictive modeling framework.

We compare the performance of the default MatMiner featurizer on MODNet in predicting the heat of formation of perovskites against implementations that incorporate additional features and substitutions with latent-space representations derived from pGNN. The Mean Absolute Error (MAE) serves as the primary evaluation metric, consistently applied alongside

a five-fold cross-validation method as described in Matbench[26]. A supporting data repository with detailed results of our work is available at https://github.com/rogeriog/Bridging-Feature-Based-Models-and-Graph-Neural-Networks_Supporting-Data.

The following outlines the specific pGNN featurizer, that are employed:

● **ℓ-OFM pGNN featurizer:** the OFM featurizer captures valence electron interactions at each atomic site by employing a weighted vector outer product of one-hot encoded valence orbitals for every atom (details in the *Supporting Information*). The structural representation is achieved by averaging all local OFMs. Utilizing the pGNN framework, we featurize the matbench_mp_gap dataset (MatBench v.0.1) comprising 106,113 structures with OFM, followed by training an autoencoder to derive a latent space representation. Various compression levels are tested to assess performance on the matbench_perovskites task. The latent OFM features are subsequently used as targets to train a GNN model that generates these features directly from the initial structures. The **CustomMEGNetFeaturizer** class in the pGNN package is implemented to retrieve these OFM-encoded features.

● **ℓ-MM pGNN featurizer :** following a similar procedure to the OFM featurizer, we encode features obtained from the default MatMiner featurizer of MODNet v.0.1.13 applied to the matbench_mp_gap dataset, resulting in 1,336 MatMiner features. Different compression levels are evaluated on both the matbench_perovskites and matbench_mp_gap tasks. The selected level of compression provides latent MatMiner features, which are then used as targets to train a GNN model that directly generates these features from the original structures. The **CustomMEGNetFeaturizer** class is also used here to retrieve these MatMiner-encoded features from the pGNN package.

● **MVL pGNN featurizers:** utilizing the **MVLFeaturizer** class from the pGNN package, we incorporate five pretrained MEGNet models provided by the Materials Virtual Lab[27].

Specifically, the models trained for the formation energy, the Fermi energy and the elastic constants K^{VRH} and G^{VRH} on the 2019.4.1 Materials Project crystals dataset, as well as the band gap regression model trained on the 2018.6.1 Materials Project crystals dataset. The default MEGNet architecture comprises MEGNet blocks followed by an MLP with two dense layers, one with 32 neurons and another with 16 neurons, before producing the target property (see **Error! Reference source not found.**). The pGNN package's modularity allows us to extract features from different layers of these pretrained models. We extract features from the MLP layers preceding the output, specifically from the 32-neuron (layer32) and 16-neuron (layer16) configurations. The extracted features (160 descriptors for layer32 and 80 descriptors for layer16) are concatenated and added to the final feature vector. These features are then evaluated on the matbench_perovskites task.

- **Adjacent pGNN featurizer:** the **AdjacentMEGNetFeaturizer** class from the pGNN package is employed to train a MEGNet model on the fly for each fold of the train-test split, using elemental embeddings transferred from the MP-crystals-2018.6.1[27] heat of formation task. This adjacent model captures task-specific data nuances, enhancing prediction accuracy. The default hyperparameters from MEGNet v.1.3.2 are utilized, as detailed in the *Supporting Information*, with the optimal 32-neuron MLP layer serving as the output for this featurizer.

Finally, the features generated with the pGNN featurizers are all tested, and their results are discussed in comparison to the pristine features and the dimensionally reduced ones. The synergic effect of including all these features is evaluated for performance on the matbench_perovskites task. Additional tests are also conducted in two tasks with comparable dataset sizes, predicting the stability and band gap of halides obtained from the OQMD database.

The pGNN package's flexibility and modular design enable us to tailor the feature extraction process. By allowing the selection of different layers from pretrained models and easy

integration of other GNN models, pGNN facilitates customized feature sets aligned with specific dataset requirements, enhancing the adaptability and performance of our predictive modeling pipeline. For detailed examples and documentation on how to implement these featurizers, please refer to pGNN's GitHub repository.

Recovering interpretability of GNN features through SHAP analysis

To identify the key features utilized by the MODNet model with GNN or latent-space features, we conducted SHAP value analysis (the definitions are given in the *Supporting Information*). We train surrogate XGBoost models to extract the most relevant GNN/latent-space features from interpretable features. Applying SHAP analysis to these models reveals correlations between key latent features and intuitive chemical descriptors, thereby restoring the model's interpretability. The SHAP analysis is implemented using the *shap* python library[28], generating SHAP summary plots with 300 samples and 500 perturbations takes less than one hour on 24 CPU cores.

XGBoost[29] is chosen as the surrogate model due to its additive and independent tree structure, which facilitates rapid parallel SHAP calculations. This choice allows us to efficiently decompose the contributions of chemical descriptors on the encoded GNN/latent-space features, training the surrogate models in minutes. All pGNN features are correlated with interpretable features derived from the same structures. For latent-space OFM, this correlation is achieved by directly comparing with original OFM features. For MVL-32 and adjacent pGNN features, we leverage other interpretable features by training XGBoost models to predict each latent feature using the original MatMiner and OFM features from the initial training dataset. These models facilitate SHAP value calculations, providing clear insights into the relationships between selected features and interpretable properties, as illustrated in Fig. 6.

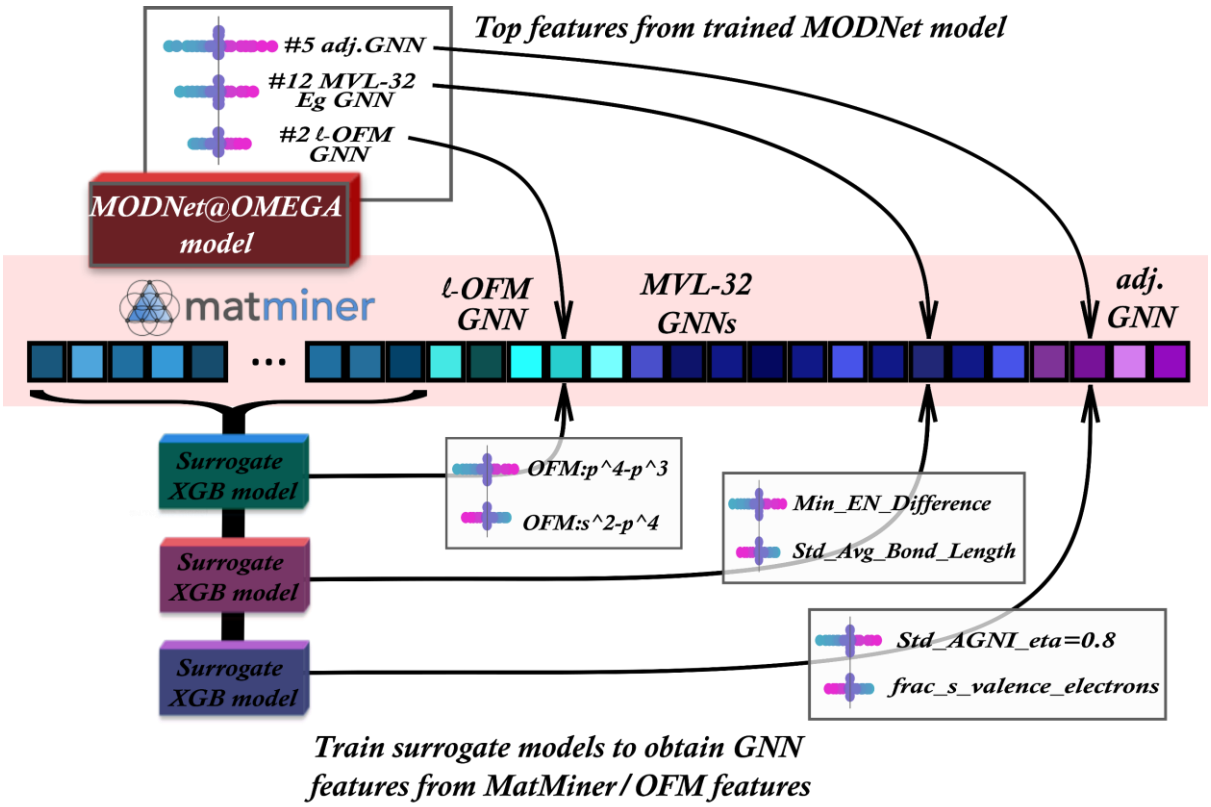

Fig. 6 | Illustration of the usage of XGBoost models to decompose the most important pGNN features from the MODNet model into intuitive chemical features via SHAP plots.

Acknowledgements

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001. Computational resources have been provided by the supercomputing facilities of the Université catholique de Louvain (CISM/UCL) and the Consortium des Équipements de Calcul Intensif en Fédération Wallonie Bruxelles (CÉCI) funded by the Fond de la Recherche Scientifique de Belgique (F.R.S.-FNRS) under convention 2.5020.11 and by the Walloon Region.

References

1. Rodrigues, J. F., Florea, L., De Oliveira, M. C. F., Diamond, D. & Oliveira, O. N. Big data and machine learning for materials science. *Discov Mater* **1**, 12 (2021).
2. Dey, A. *et al.* State of the Art and Prospects for Halide Perovskite Nanocrystals. *ACS Nano* **15**, 10775–10981 (2021).

3. Guo, K., Yang, Z., Yu, C.-H. & Buehler, M. J. Artificial intelligence and machine learning in design of mechanical materials. *Mater. Horiz.* **8**, 1153–1172 (2021).
4. De Breuck, P.-P., Hautier, G. & Rignanese, G. M. Materials property prediction for limited datasets enabled by feature selection and joint learning with MODNet. *npj Computational Materials* **7**, 1–8 (2021).
5. Zhang, B., Zhou, M., Wu, J. & Gao, F. Predicting the Materials Properties Using a 3D Graph Neural Network With Invariant Representation. *IEEE Access* **10**, 62440–62449 (2022).
6. Tawfik, S. A. & Russo, S. P. Naturally-meaningful and efficient descriptors: machine learning of material properties based on robust one-shot ab initio descriptors. *J Cheminform* **14**, 78 (2022).
7. Ward, L. *et al.* Matminer: An open source toolkit for materials data mining. *Computational Materials Science* **152**, 60–69 (2018).
8. Pretto, T., Baum, F., Gouvêa, R. A., Brolo, A. G. & Santos, M. J. L. Optimizing the Synthesis Parameters of Double Perovskites with Machine Learning Using a Multioutput Regression Model. *J. Phys. Chem. C* **128**, 7041–7052 (2024).
9. Liu, J. *et al.* Toward Excellence of Electrocatalyst Design by Emerging Descriptor-Oriented Machine Learning. *Adv Funct Materials* **32**, 2110748 (2022).
10. Kim, G.-H., Lee, C., Kim, K. & Ko, D.-H. Novel structural feature-descriptor platform for machine learning to accelerate the development of organic photovoltaics. *Nano Energy* **106**, 108108 (2023).
11. Li, S. *et al.* Encoding the atomic structure for machine learning in materials science. *WIREs Comput Mol Sci* **12**, e1558 (2022).
12. Dunn, A. MatBench Leaderboard. https://matbench.materialsproject.org/ (2024).

13. Xie, T. & Grossman, J. C. Crystal Graph Convolutional Neural Networks for an Accurate and Interpretable Prediction of Material Properties. *Phys. Rev. Lett.* **120**, 145301 (2018).
14. Lam Pham, T. *et al.* Machine learning reveals orbital interaction in materials. *Science and Technology of Advanced Materials* **18**, 756–765 (2017).
15. Bartók, A. P., Kondor, R. & Csányi, G. On representing chemical environments. *Phys. Rev. B* **87**, 184115 (2013).
16. Oviedo, F., Ferres, J. L., Buonassisi, T. & Butler, K. T. Interpretable and Explainable Machine Learning for Materials Science and Chemistry. *Acc. Mater. Res.* **3**, 597–607 (2022).
17. Pilania, G. Machine learning in materials science: From explainable predictions to autonomous design. *Computational Materials Science* **193**, 110360 (2021).
18. Abolhasani, M. & Kumacheva, E. The rise of self-driving labs in chemical and materials sciences. *Nat. Synth* **2**, 483–492 (2023).
19. matbench_v0.1: MegNet (kgcnn v2.1.0) - MatBench. *Materialsproject.org* https://matbench.materialsproject.org/Full%20Benchmark%20Data/matbench_v0.1_MegNet_kgcnn_v2.1.0/ (2020).
20. Schmidt, J., Pettersson, L., Verdozzi, C., Botti, S. & Marques, M. A. L. Crystal graph attention networks for the prediction of stable materials. *Sci. Adv.* **7**, eabi7948 (2021).
21. Sa, R., Luo, B., Ma, Z. & Liu, D. The effect of the A-site cation on the stability and physical properties of vacancy-ordered double perovskites A_2PtI_6 (A = Tl, K, Rb, and Cs). *Journal of Solid State Chemistry* **305**, 122714 (2022).
22. Shen, J. *et al.* Reflections on one million compounds in the open quantum materials database (OQMD). *J. Phys. Mater.* **5**, 031001 (2022).
23. Bartel, C. J. Review of computational approaches to predict the thermodynamic stability of inorganic solids. *J Mater Sci* **57**, 10475–10498 (2022).

24. Riebesell, J. Matbench Discovery. https://janosh.github.io/matbench-discovery (2024).
25. Choudhary, K. & DeCost, B. Atomistic Line Graph Neural Network for improved materials property predictions. *npj Comput Mater* **7**, 185 (2021).
26. Dunn, A., Wang, Q., Ganose, A., Dopp, D. & Jain, A. Benchmarking materials property prediction methods: the Matbench test set and Automatminer reference algorithm. *npj Computational Materials* **6**, 1–10 (2020).
27. Chen, C., Ye, W., Zuo, Y., Zheng, C. & Ong, S. P. Graph Networks as a Universal Machine Learning Framework for Molecules and Crystals. *Chem. Mater.* **31**, 3564–3572 (2019).
28. Lundberg, S. M. & Lee, S.-I. A unified approach to interpreting model predictions. in *Advances in neural information processing systems* (eds. Guyon, I. et al.) vol. 30 (Curran Associates, Inc., 2017).
29. Chen, T. & Guestrin, C. XGBoost: A Scalable Tree Boosting System. in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 785–794 (ACM, San Francisco California USA, 2016). doi:10.1145/2939672.2939785.