Verbindung

A Physics-Aware Multi-Task Learning Framework for Coupled Prediction of Structural and Electronic Properties in Materials

[1]Omprakash Barapatre, [2]Shubham Sanskar Routray, [3]Shruti Patel, [4]Tottaramudi Jhansi Victoriya

[1]Assistant Professor, [2,3,4]Student

[1]Department of Computer Science and Engineering,

[1]Bhilai Institute of Technology,Raipur,India

[1]omprakashbarapatre@bitraipur.ac.in,[2]shubhamroutray4@gmail.com, [3]shrutipatel5290@gmail.com,[4]jhansivictoriyat@gmail.com

Abstract **:** Accurate prediction of physical properties, such as band gap, is crucial in the design and characterization of novel materials with unique functionalities. Although DFT-based calculations are considered reliable and have been used extensively in the field, they are time-consuming and computationally expensive; therefore, they cannot be applied on a broad scale in materials screening. This problem can be solved using a hybrid machine learning methodology aimed at predicting the properties of new materials with a focus on their reliability.

For predicting the band gap property, different regression models are proposed in this work, including Random Forest, XGBoost, ExtraTrees, CatBoost, Multi-Layer Perceptron (MLP), and stacking. The training set is prepared by means of an elaborate feature engineering pipeline, which utilizes multiple descriptors, including elemental, stoichiometric, valence orbital descriptors, and ion-related features. In addition, we employ data augmentation methods that enhance synthetic data alongside real samples to improve the performance of the model.

In this study, we also propose a new method for evaluating the reliability of predictions based on the combined use of a model's uncertainty estimates along with out-of-distribution (OOD) score, estimated with Isolation Forest and nearest-neighbor distance algorithms.

From experimental results, it is evident that the proposed framework demonstrates high predictive ability. As for the performance among all considered machine learning models, it can be observed that the XGBoost model outperforms others with R^2 of 0.912, MAE of 0.250 eV, and RMSE of 0.349 eV. At the same time, the second best-performing model is Random Forest with R^2 of 0.905, MAE of 0.248 eV, and RMSE of 0.361 eV. Stacking ensemble achieves an R^2 of 0.901, and ExtraTrees ($R^2 = 0.878$), MLP ($R^2 = 0.869$), and CatBoost ($R^2 = 0.861$) demonstrate comparative results. At the same time, GNN model demonstrates weak performance with negative R^2, demonstrating shortcomings of the structure-based approach when data availability is insufficient.

Conclusively, the results of experiments demonstrated the efficiency of hybrid ensemble learning approaches to property prediction, and highlighting it.
Index Terms —Materials Informatics, Band Gap Prediction, Machine Learning, Graph Neural Networks, Ensemble Learning, Uncertainty Quantification, Out-of-Distribution Detection, Materials Discovery.

I. INTRODUCTION

Material discovery entails the investigation and integration of novel materials exhibiting specific electronic characteristics. The band gap of a material is important for figuring out if it can be used in

semiconductor, photovoltaic, optoelectronic, and energy storage technologies[1]. Correctly estimating the band gap lets you properly evaluate potential materials, which speeds up the process of finding new materials.

Band gap prediction conventionally relies on theoretical methods such as density functional theory (DFT). While accurate, DFT computations demand substantial computational power and time, especially when employed in screening large material datasets[2],[3]. This has led to the adoption of machine learning approaches that can learn the complex relationship between the chemical structure and properties of materials from existing data sources and make fast predictions[1],[6].

Over the past few years, different machine learning models such as random forest, gradient boosting, neural networks, among others, have been successfully used for predicting band gaps[1],[6]. Another algorithm gaining attention because of its ability to incorporate structural information in materials is the Graph Neural Networks (GNN)[3], [4]. Despite all these advancements, most techniques focus on improving predictive performance without considering another important aspect: the uncertainty of predictions. The reliability of any prediction in practical scenarios is not only about the ability to predict but also about being able to measure the uncertainty and the confidence level in the prediction[7],[17].

The main challenge that faces prediction in materials science through machine learning approaches is out-of-distribution (OOD) samples[8]; the difference between the test and training data is high. As a result, these models make predictions that seem convincing but are completely wrong. Ensemble models make multiple predictions and fail to use their variations in estimating uncertainties.

To tackle these issues, this study presents a hybrid machine learning framework for predicting band gaps that combines ensemble learning with a reliability-focused decision system. The suggested approach integrates Random Forest and XGBoost models in a stacking framework, accompanied by a meta-learning layer to improve predictive accuracy. A thorough feature engineering pipeline utilizes composition-based descriptors, encompassing elemental properties, stoichiometric attributes, valence orbital traits, and ion-related features.

Along with enhancing predictive accuracy, this study presents a new reliability estimation method that integrates model uncertainty with out-of-distribution detection. Uncertainty is measured by the standard deviation of ensemble forecasts, while out-of-distribution detection is conducted using Isolation Forest and nearest-neighbor distance metrics. These elements are combined into a cohesive reliability score, allowing the system to differentiate between dependable and undependable forecasts.

The key contributions of this research can be outlined as follows[18]:

- Creation of a hybrid ensemble learning model for precise band gap forecasting.
- Incorporation of a meta-learning framework to improve forecasting accuracy.
- Implementation of a prediction system that considers reliability, integrating uncertainty estimation and out-of-distribution detection.
- Utilization of both authentic and artificially created data to enhance model generalization.

The suggested method offers a scalable and effective solution for predicting material properties, focusing on both reliability and precision. This renders it especially appropriate for practical materials discovery tasks where dependable decision-making is essential.

II. LITERATURE REVIEW

Recent progress in machine learning has notably sped up the forecasting and identification of material properties in various material systems, such as perovskites, crystalline solids, and intricate alloys. Combining data-driven approaches with materials science has facilitated effective investigation of structure–property relationships, lessening dependence on costly first-principles methods[2],[3].

Conventional machine learning methods have shown excellent results in predicting material properties when paired with efficient feature engineering. Chen et al. (2024) employed a dataset of perovskite structures computed through density functional theory (DFT) to forecast band gap values using

descriptors like atomic radius, electronegativity, and structural characteristics. Their research indicated that models like Random Forest, Gradient Boosting, and Deep Neural Networks can attain significant predictive accuracy, emphasizing the necessity of well-crafted compositional and structural attributes.

Graph-based deep learning methods have recently emerged as a potent framework for modeling materials on the atomic level. Selvaraj (2024) suggested modeling crystal structures as graphs, with atoms represented as nodes and bonds as edges. By employing Graph Neural Networks (GNNs) and message passing techniques, the model successfully captures interatomic interactions, leading to enhanced predictions of electronic and structural characteristics. A thorough analysis by Shi et al. (2024) reinforces the importance of GNNs in materials science, discussing atomistic graph representations, strategies for feature extraction, and learning the relationship between structure and property. The research highlights that message passing neural networks excel at capturing intricate atomic interactions.

Expanding on these principles, Zhou et al. (2025) presented a deep graph network model that includes hierarchical message passing and bond-angle characteristics. By combining atom-level, edge-level, and geometric data, their method shows enhanced prediction accuracy via physics-informed modeling. In a similar manner, Du et al. (2024) introduced the Crystal Transformer Graph Neural Network (CTGNN), integrating graph neural networks with transformer structures to grasp local atomic interactions and global structural dependencies, thereby addressing the shortcomings of traditional GNN models[15].

Apart from crystalline structures, machine learning has been extensively utilized in intricate alloy systems[5],[6]. Golbabaei et al. (2025) examined the use of machine learning in high-entropy alloys, highlighting its significance in feature engineering, predicting properties, identifying phases, and discovering materials. Rahman et al. (2025) expanded this viewpoint by presenting an extensive review of machine learning methodologies for various alloy systems, covering supervised, unsupervised, reinforcement learning, and physics-informed techniques for analyzing phase stability and optimizing materials.

A significant challenge in materials informatics is the lack of data. Li et al. (2025) tackled this problem with the MatWheel framework, which integrates synthetic data creation with actual datasets by utilizing generative models and semi-supervised learning[13]. Their research shows that enhancing sparse experimental data with synthetic examples can greatly enhance model generalization and predictive accuracy[14].

Overview of Findings from Research:

- Traditional machine learning models yield positive results when paired with robust feature engineering.
- Graph Neural Networks offer a robust framework for representing interactions at the atomic scale.
- Physics-informed attributes (e.g., bond angles, distances) greatly improve model performance.
- Hybrid architectures (e.g., GNN combined with Transformer) enhance representation learning.
- Machine learning finds extensive application across various classes of materials, such as alloys and crystalline solids.
- Data shortage continues to be a significant problem, with artificial data appearing as a hopeful remedy.

Recognized Research Gaps:

- We identified that the majority of research depends on single-model methods, missing hybrid integration.
- As minimal research merges composition-driven descriptors with structure-oriented graph learning.
- We found uncertainty quantification and reliability evaluation are frequently overlooked.
- There were Inadequate methods for identifying out-of-distribution (OOD) items.

- Also Data scarcity is addressed only in part, with minimal incorporation into comprehensive predictive systems.
- We found the limited application of active learning frameworks for ongoing model enhancement.
- Apart that rarely do systems offer tools for interactive and real-time exploration of materials.

III. METHODOLOGY

3.1 Overview

The proposed framework is a hybrid AI-driven materials discovery system that integrates:

- Composition-based feature engineering
- Structure-based graph learning
- Ensemble machine learning models
- Uncertainty-aware prediction
- Out-of-distribution (OOD) detection

The overall pipeline is illustrated as:

Input Composition → Feature Extraction → ML Models + GNN → Ensemble Prediction → Reliability Estimation

3.2 Data Preparation

The dataset consists of:

- Real experimental/DFT data
- Synthetic compositions generated programmatically

Let the dataset be:

$$D = \{(x_i, y_i)\}_{i=1}^{N}$$

where:

- x_i = material composition
- y_i = target property (band gap)

Synthetic data is generated to improve generalization:

$$y_{synth} = \alpha . \underline{Z} + \beta . \triangle \chi + \epsilon$$

where:

- $\underline{Z}$ = average atomic number
- $\triangle \chi$ = electronegativity difference
- $\epsilon \sim N(0, \sigma^2)$

3.3 Feature Engineering

Material compositions are converted into numerical feature vectors using matminer descriptors.

Feature vector:

$$f = [f_1, f_2, \ldots\ldots, f_n]$$

Additional physics-based features:

- Average atomic number:

$$\underline{Z} = \frac{1}{N}\sum_{i=1}^{N} Z_i$$

Electronegativity difference:

$$\triangle \chi = max(\chi_i) = min(\chi_i)$$

- Atomic radius variance:

$$\sigma_r^2 = \frac{1}{N}\sum (r_i - \underline{r})^2$$

3.4 Data Preprocessing

Missing Value Imputation

$$X' = Imputer(X)$$

Feature Selection (F-test)

$$F = \frac{Variance\ between\ groups}{variance\ within\ groups}$$

Top k features are selected.

Dimensionality Reduction (PCA)

$$Z = XW$$

where:

- W = principal components

Normalization

$$X_{scaled} = \frac{X - \mu}{\sigma}$$

3.5 Machine Learning Models

Random Forest

$\hat{y}_{XGB} = \frac{!}{T}\sum_{t=1}^{T} h_i(x)$

where:

- h_i = decision tree

XGBoost

$$\hat{y}_{XGB} = \sum_{k=1}^{K} f_k(x)$$

with objective:

$$L = \sum l(y_i, \hat{y}_i) + \sum \Omega(f_x)$$

Stacking Model

$$\hat{y}_{stack} = g(\hat{y}_{RF}, \hat{y}_{XGB})$$

where:

- g = meta-learner (Linear Regression)

3.6 Graph Neural Network (GNN)

Materials are represented as graphs:

G=(V,E)

where:

- V = atoms (nodes)
- E = bonds (edges)

Node features:

$$x_i = [Z, \chi, m, row, group, oxidation]$$

Edge features:

$$e_{ij} = [d_{ij}, w_{ij}]$$

Message Passing

$$h_i = \sigma(\sum_{j \in N(i)} \alpha_{ij} W h_j^{(l)})$$

where:

- α_{ij} = attention weight
- W = learnable weight matrix

Attention Mechanism

$$\alpha_{ij} = \frac{exp(LeakyReLU(a^T[Wh_i| \quad |Wh_j]))}{\sum_{k \in N(i)} \quad exp(\dots)}$$

Graph Pooling

$$h_G = \frac{1}{|V|} \sum_{i \in V} \quad h_i$$

Output

$$\hat{y}_{GNN} = f(h_G)$$

3.7 Ensemble Prediction

Final prediction is computed as:

$$\mu = \frac{1}{M} \sum_{m=1}^{M} \hat{y}_m$$

where:

- M = 3 or 4 models

3.8 Uncertainty Estimation

Standard deviation of predictions:

$$\sigma = \sqrt{\frac{1}{M} \sum_{m=1}^{M} (\hat{y}_m - \mu)^2}$$

3.9 Out-of-Distribution (OOD) Detection

Isolation Forest Score:

$$D_{iso} = decision\_function(x)$$

Distance-based OOD:

$$D_{knn} = \frac{1}{k}\sum_{i=1}^{k} d(x, x_i)$$

Final OOD score:

$$OOD = \lambda_1(1 - D_{iso}) + \lambda_2 D_{knn}$$

3.10 Reliability Score

$$R = e^{-\sigma - OOD}$$

Decision rule:

$$Decision = \begin{cases} Accept\ R > 0.5 \\ Reject\ otherwise \end{cases}$$

3.11 Active Learning

New samples are selected based on high uncertainty:

$$x^* = arg\ max\ \sigma(x)$$

Model is updated:

$$D \leftarrow D \cup \{(x^*, \hat{y})\}$$

3.12 Visualization Module

- 3D crystal structures (py3Dmol)
- Graph visualization (Plotly)
- Attention-based bond importance

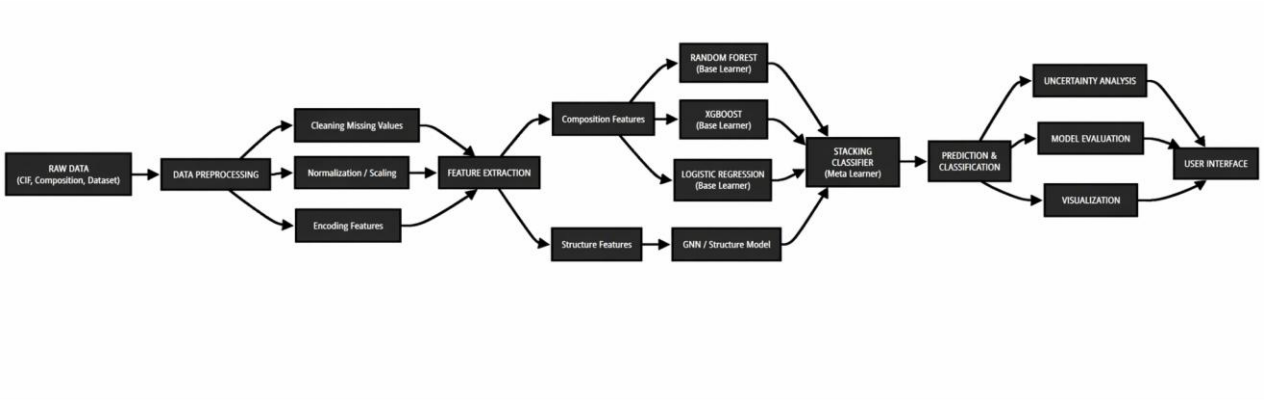

Fig 1.Flowchart

IV. MACHINE LEARNING MODELS USED

This research utilized a hybrid ensemble of deep learning and machine learning models to precisely forecast the band gap of materials based on their structural features and chemical composition. Conventional machine learning models, such as Random Forest and Extreme Gradient Boosting (XGBoost), were employed as primary learners because of their excellent ability to manage nonlinear relationships and high-dimensional feature spaces[1],[6]. Random Forest, as an ensemble of decision trees, successfully identified intricate relationships among elemental traits, while XGBoost offered superior optimization via gradient boosting and regularization, minimizing overfitting and improving generalization[1][6].

To enhance predictive performance, a stacking ensemble method was utilized, in which the predictions from separate models were merged using a linear regression-based meta-learner. This method utilized the advantages of every base model and generated more reliable and consistent predictions. Moreover, a Graph Neural Network (GNN) utilizing Graph Attention Networks (GAT) was integrated to capture atomic interactions at a granular level by modeling crystal structures as graphs, with atoms represented as nodes and bonds as edges[3],[4]. The attention mechanism enabled the model to dynamically grasp the significance of various atomic interactions.

Model	R^2 Score	MAE(eV)	RMSE(eV)
XGBoost	0.912	0.250	0.349
Random Forest	0.905	0.248	0.361
Stacking	0.901	0.250	0.369
ExtraTrees	0.878	0.264	0.410
MLP	0.869	0.299	0.425
CatBoost	0.861	0.323	0.438
Mean Ensemble	0.800	0.395	0.525
GNN	0.451	1.040	1.414

Fig 2.Comparison Table.

A hybrid meta-model was subsequently created to combine forecasts from both classical machine learning models and the graph-oriented deep learning model. This integration allowed the system to leverage both compositional characteristics and structural data, leading to enhanced accuracy and dependability. Additionally, an anomaly detection system based on Isolation Forest and nearest-neighbor distance evaluation were employed to evaluate out-of-distribution samples and gauge the reliability of predictions[8]. This thorough integration of models guaranteed that the system not only attained high prediction accuracy but also offered uncertainty assessment and resilience for practical materials discovery applications.

V. RESULTS AND DISCUSSION

In order to thoroughly assess the effectiveness of the suggested hybrid framework, various machine learning and deep learning models were examined, such as tree-based techniques, boosting methods, neural networks, and graph-oriented strategies. The assessment utilized standard regression metrics

including R^2 score, Mean Absolute Error (MAE), and Root Mean Square Error (RMSE) on a dataset of around 29,379 material compositions. This dataset was created by combining actual experimental data with synthetically produced samples to enhance generalization[13]. The entire system workflow, depicted in Figure 2, showcases the combination of feature engineering based on composition, optional processing of structural graphs, ensemble modeling, and a prediction mechanism that considers reliability.

Out of all the assessed models, XGBoost was identified as the top-performing algorithm, obtaining an R^2 score close to 0.9117 and the minimum RMSE of 0.3488, demonstrating excellent predictive accuracy and robust generalization ability. The Random Forest model demonstrated a strong performance with an R^2 score of 0.9052 and the minimum MAE of 0.2479, emphasizing its reliability and uniformity throughout the dataset. The stacking ensemble method reached an R^2 score of 0.9013, showing that integrating various models enhances durability and prediction consistency, even if it does not necessarily exceed the top-performing individual model. These findings unmistakably show that ensemble tree-based models are extremely efficient for structured, composition-focused materials data.

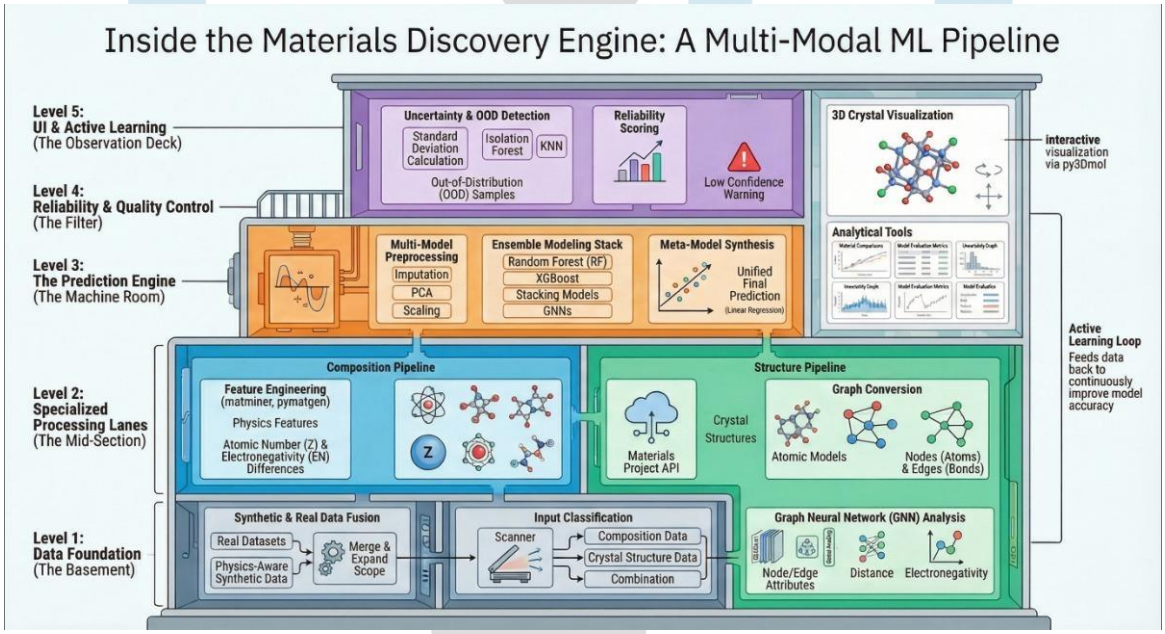

Fig 3.Model Pipeline

To provide a broader comparison, additional models were evaluated. The ExtraTrees model achieved an R^2 score of 0.8777, showing strong performance but slightly reduced generalization compared to Random Forest. CatBoost obtained an R^2 score of 0.8609, suggesting that its advantages are less pronounced in purely numerical feature spaces. The Multi-Layer Perceptron (MLP) achieved an R^2 score of 0.8688, demonstrating its ability to capture nonlinear relationships; however, it underperformed compared to boosting methods, likely due to the limited size and structured nature of the dataset. The mean ensemble approach, which averages predictions from multiple models, resulted in a significantly lower R^2 score of 0.7996, indicating that simple averaging fails to capture complex relationships effectively without learned weighting mechanisms.

In contrast, the Graph Neural Network (GNN) model exhibited poor performance, with negative R^2 values and significantly higher error metrics. This indicates that the model failed to learn meaningful structural representations under the given conditions. The underperformance of the GNN can be attributed to limited availability of high-quality crystal structure data, insufficient training, and relatively simple graph representations. This observation highlights a key insight of this study: although

graph-based deep learning models are theoretically powerfu[3],[4]l, they require large-scale, high-quality structural datasets to outperform traditional machine learning methods.

The predictive capability of the proposed framework is further validated through the true vs. predicted scatter plot shown in Figure 3, where a strong linear correlation is observed between actual and predicted band gap values. Most data points are closely aligned along the diagonal, indicating that the model effectively captures the underlying distribution of the data. Minor deviations are observed at higher band gap values, suggesting slightly increased prediction difficulty in extreme regions; however, the overall trend confirms strong generalization performance.

Fig 4. True vs Predicted

The practical effectiveness of the system is demonstrated through the interactive dashboard outputs. In the batch prediction results (Figure 4), materials such as Fe_2O_3, Al_2O_3, and SiO_2 are evaluated, providing predicted band gap values along with reliability scores. It is observed that Fe_2O_3 exhibits a significantly lower reliability score, indicating higher uncertainty in its prediction. This observation is further supported in the comparison analysis (Figure 5), where Fe_2O_3 shows higher uncertainty and out-of-distribution (OOD) scores, suggesting that it lies farther from the training data distribution. In contrast, Al_2O_3 and SiO_2 exhibit lower uncertainty and OOD values, resulting in more reliable and trustworthy predictions.

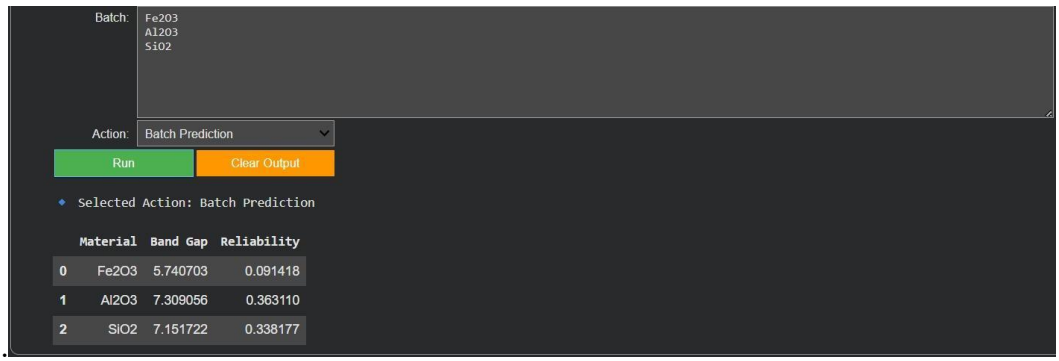

Fig 5.Batch Prediction

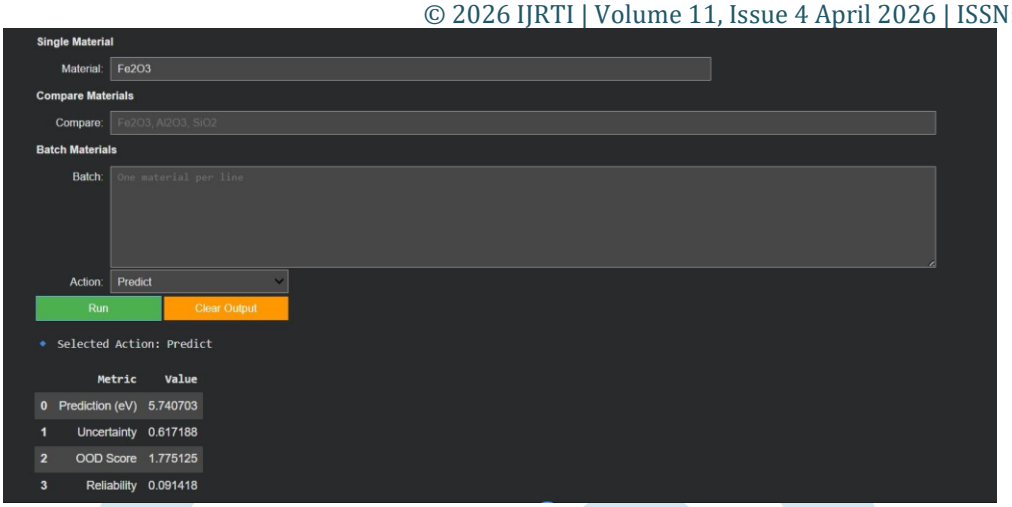

Fig 6.Comparison Analysis

The uncertainty analysis presented in Figure 6 further reinforces these findings, where Fe_2O_3 demonstrates the highest prediction variance, while Al_2O_3 shows the lowest. This confirms the effectiveness of the uncertainty estimation mechanism in identifying less confident predictions[7].

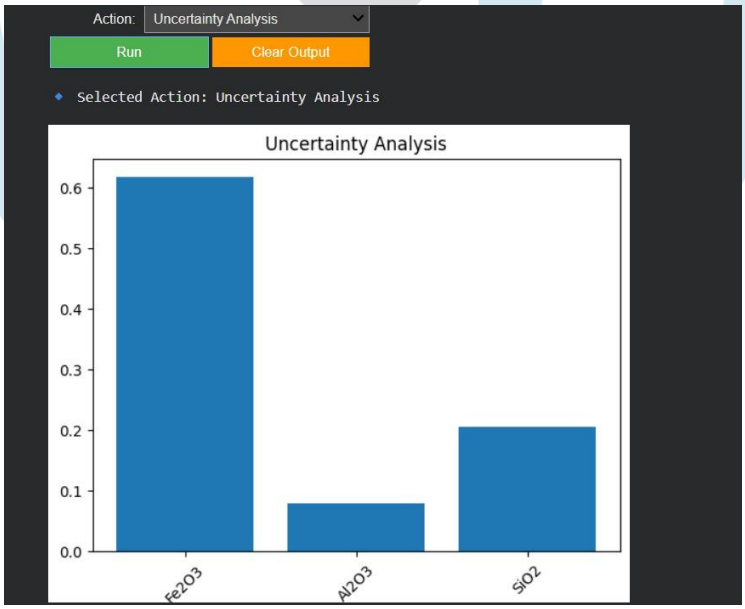

Fig 7. Uncertainty Analysis

A key contribution of this work is the reliability estimation framework, which integrates prediction variance across models, anomaly detection using Isolation Forest, and nearest-neighbor distance analysis. By combining these components, the system computes a reliability score that reflects both model confidence and how closely the input aligns with the training data distribution. This enables the identification of uncertain or out-of-distribution predictions, significantly enhancing decision-making in materials discovery.

Overall, the results demonstrate that the proposed hybrid framework achieves strong predictive performance while addressing a critical limitation of traditional models through reliability-aware prediction. The findings confirm that tree-based ensemble models outperform neural and graph-based approaches for composition-driven datasets, while deep learning methods require more extensive and structured data to realize their full potential. The integration of ensemble learning, uncertainty quantification, and interactive visualization makes the proposed system a robust, scalable, and practical solution for real-world materials informatics applications.

VI. CONCLUSION

This research established a thorough machine learning framework aimed at predicting the band gap of materials by utilizing compositional and structural data. Various models such as Random Forest, XGBoost, ExtraTrees, CatBoost, Multi-Layer Perceptron (MLP), and Graph Neural Networks (GNN) were developed and assessed. Among these, XGBoost delivered the top performance with the highest R^2 score and the least prediction error, closely followed by Random Forest and stacking ensemble models. The findings indicate that ensemble methods based on trees are very efficient in managing structured, high-dimensional compositional data.

The incorporation of ensemble learning methods boosted prediction stability, while adding uncertainty estimation and out-of-distribution detection increased the system's dependability. Despite the significant theoretical promise of Graph Neural Networks in capturing atomic-level interactions, their effectiveness in this research was hindered by inadequate and inconsistent structural data. This emphasizes the significance of data quality and representation within materials informatics[14].

The suggested framework offers a precise, dependable, and scalable method for predicting material properties[15],[16]. It can act as an effective aid for speeding up materials discovery by lessening the reliance on expensive experimental and computational techniques.

Future Scope

Despite the promising results, several opportunities exist for further improvement and expansion of this work:

- Integration of Premium Structural Information:Future efforts may center on merging extensive, high-quality crystal structure datasets to enhance the capabilities of Graph Neural Networks and maximize the benefits of structure-based learning.
- Sophisticated Graph Structures:Utilizing advanced models like transformer-based graph networks (e.g., Crystal Transformer GNN) can improve the capability to grasp both local and global atomic interactions.
- Tuning Hyperparameters:Techniques for automated tuning, like Bayesian optimization or evolutionary algorithms, can be utilized to enhance model performance further.
- Extension to Multi-Property Forecasting:The framework is capable of being enhanced to concurrently predict various material properties, including formation energy, stability, and conductivity.
- Enhanced Material Discovery Workflow:Improving the discovery module using more advanced search techniques (such as reinforcement learning or genetic algorithms) could result in the recognition of new high-performance materials.
- Integration with Experimental Verification:Working with experimental studies to confirm predicted materials can connect computational forecasts with practical applications.
- Dealing with Limited Data:Methods like synthetic data creation, transfer learning, and semi-supervised learning can be investigated to address the challenges posed by small datasets.

References

1. Journal of Materials Chemistry C. (2024). Machine learning methods for predicting perovskite and semiconductor material properties. Journal of Materials Chemistry C.
2. npj Computational Materials. (2024). Machine learning prediction of electronic properties of inorganic materials. npj Computational Materials.
3. Physical Review Materials. (2024). Crystal graph neural networks for accurate crystal property prediction. Physical Review Materials.
4. Computational Materials Science. (2024). Deep message passing neural networks for crystal material property prediction. Computational Materials Science.
5. Advanced Engineering Materials. (2025). Machine learning prediction of mechanical properties in

high-entropy alloys. Advanced Engineering Materials.
6. Materials & Design. (2024). Data-driven prediction of mechanical and structural properties of alloys using machine learning. Materials & Design.
7. Nature Reviews Physics. (2024). Uncertainty quantification in machine learning models for physical sciences. Nature Reviews Physics.
8. arXiv. (2023). Identification of high-reliability regions in machine learning predictions for materials science. arXiv Preprint. https://arxiv.org
9. Materials Project. (2024). Materials Project database. https://materialsproject.org
10. Open Quantum Materials Database (OQMD). (2024). Open Quantum Materials Database. http://oqmd.org
11. Matter. (2024). Multi-task learning for simultaneous prediction of thermodynamic and electronic properties. Matter.
12. Digital Discovery. (2024). Physics-informed graph neural networks for state-of-the-art dielectric property prediction. Digital Discovery.
13. Advanced Functional Materials. (2025). Transfer learning and multi-fidelity modeling for crystal structure-property mapping. Advanced Functional Materials.
14. Scientific Data. (2024). An integrated database of structural and electronic descriptors for crystalline materials. Scientific Data.
15. Journal of Cheminformatics. (2024). Attention-based transformers for capturing long-range atomic interactions in complex crystals. Journal of Cheminformatics.
16. Acta Materialia. (2024). Machine learning-driven discovery of high-entropy ceramics with tailored electronic properties. Acta Materialia.
17. AIMS Materials Science. (2025). Evaluating the reliability of deep learning models in predicting semiconductor band gaps. AIMS Materials Science.
18. G. Liu, K. Xu, J. Zhu, Y. Ge, and X. Chen, “A local-global transformer-based model for person re-identification,” *PLoS One*, vol. 20, no. 11, e0335848, 2025. doi: 10.1371/journal.pone.0335848:https://pmc.ncbi.nlm.nih.gov/articles/PMC12614614