Graph Neural Networks for Predicting Mechanical Response in Defective Graphene

1. Project Overview

This project focuses on predicting mechanical properties of materials, specifically **Young's modulus** and **fracture strain**, using a graph-based deep learning model called **GINE (Graph Isomorphism Network with Edge features)**.

- **Objective:** Accurately predict mechanical properties from material structure.
- **Motivation:** These properties are critical in material design and engineering applications.

---

2. Dataset Description

- **Source:**Synthetic Dataset.
- **Size:** 1000.
- **Features:** Molecular graph representations, edge attributes like bond lengths.
- **Targets:** Young's modulus (`young_modulus`), fracture strain (`fracture_strain`).
- **Splits:**
  - Training: 70%
  - Validation: 15%
  - Test: 15%

Preprocessing Steps:

- Target values scaled using `StandardScaler`.
- Graph construction: atoms as nodes, bonds as edges with edge features.
- Data loaders created for batch training.

---

3. Model Architecture

- **Model:** GINE Regressor with 3 layers.
- **Input channels:** 4
- **Hidden channels:** 512
- **Output channels:** 2 (Young's modulus & fracture strain)

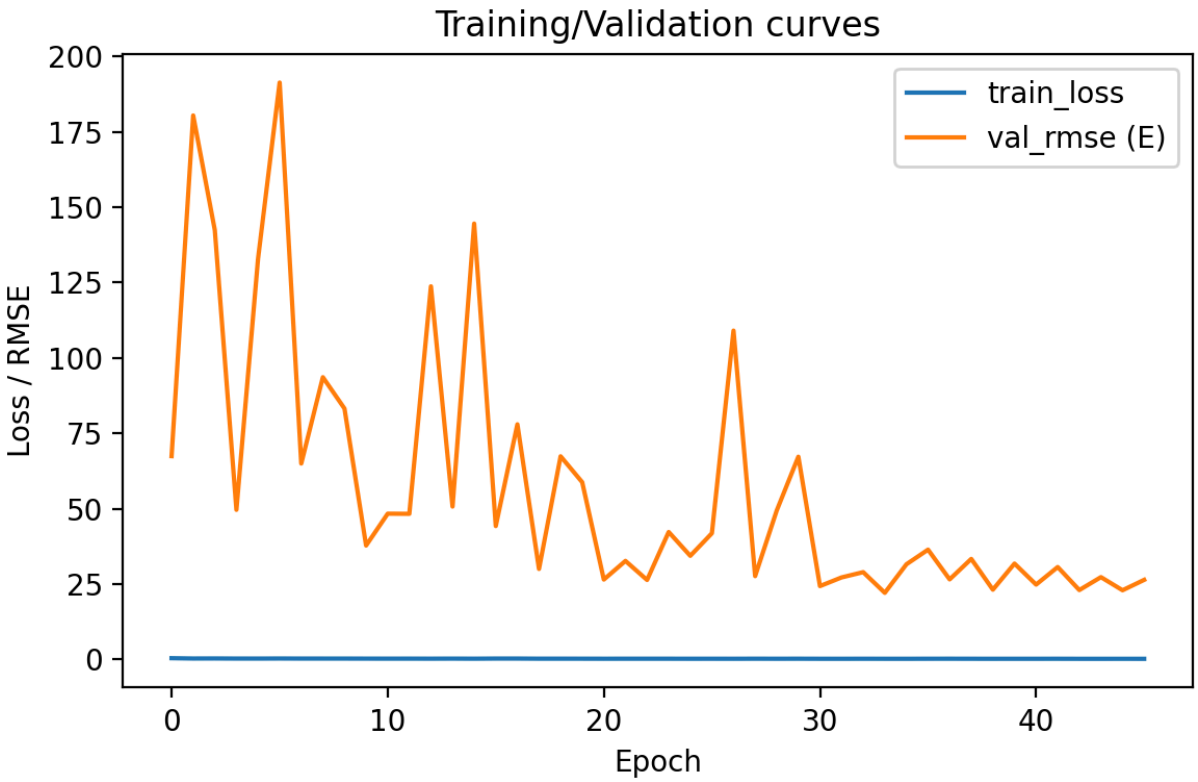
- **Dropout:** 0.1
- **Edge features:** Bond lengths
- **Batch Normalization:** Applied after hidden layers

---

4. Training Setup

- **Epochs:** 120 (Early stopping triggered at epoch 45)
- **Batch size:** 16
- **Optimizer:** Adam
- **Learning rate:** 1e-3
- **Loss Function:** Mean Squared Error (MSE)
- **Scheduler:** Reduces LR on plateau of validation RMSE
- **Early Stopping:** Patience = 12

Training Curve Placeholder:



5. Evaluation Metrics

Metrics used for regression:

- **MSE:** Mean Squared Error
- **MAE:** Mean Absolute Error
- **RMSE:** Root Mean Squared Error
- **R^2:** Coefficient of Determination

---

6. Results

Validation Performance

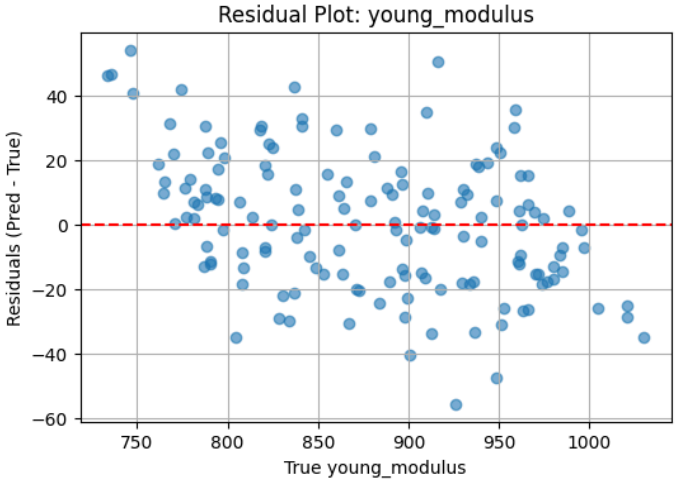
- Best **Validation RMSE (Young's modulus):** 22.0687

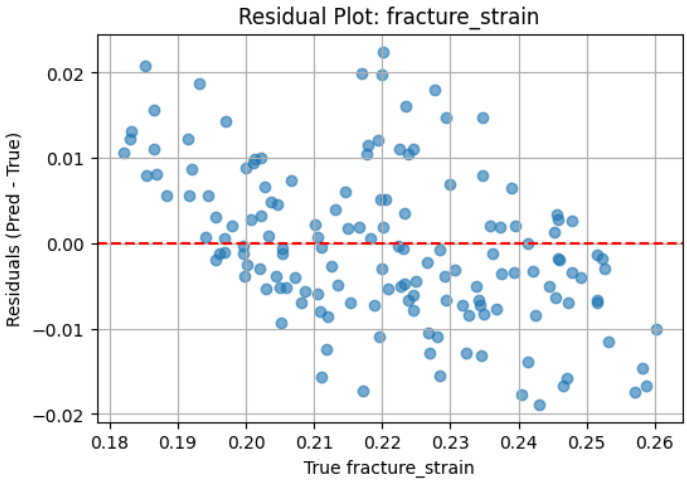
Test Set Metrics

Target	MSE	MAE	RMSE	R^2
Young's Modulus	444.9727	17.0751	21.0944	0.9180
Fracture Strain	0.000080	0.007203	0.008964	0.7891

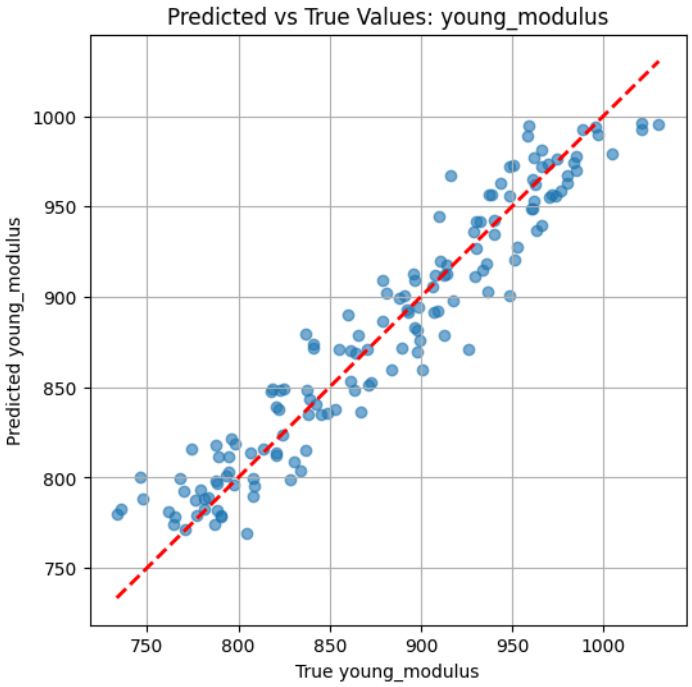
Prediction Plots Placeholders:

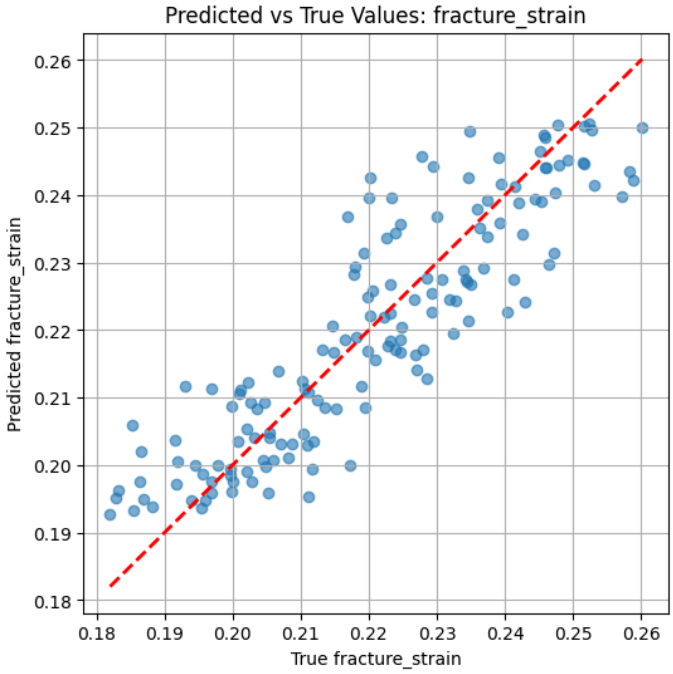
- Residual/error distribution plots :





- Scatter plot: True vs Predicted Young's modulus





---

7. Analysis & Discussion

- **Young's modulus** predictions are highly accurate (**R^2 = 0.918**).
- **Fracture strain** predictions are moderately accurate (**R^2 = 0.789**), suggesting more data or features could improve results.

- **Observations:**

    - Some epochs show high validation RMSE; model benefits from early stopping.
    - Errors for fracture strain are larger in some extreme values.

Potential Improvements:

- Add more informative edge/node features.
- Experiment with deeper GNN layers or attention mechanisms.
- Include data augmentation or additional molecular properties.

---

8. Conclusion

The GINE model successfully predicts material properties from graph representations.

- Achieved high accuracy for Young’s modulus and reasonable performance for fracture strain.
- Framework can be extended with more features or larger datasets for improved predictions.
