SemiMat: A Semi-Supervised Toolkit for Data-Scarce Materials Property Prediction

Wentao Li [*1] **Yizhe Chen** [*1] **Xiaonan Wang** [1]

[*]Equal contribution [1]*Tsinghua University, Beijing, China*. Correspondence to: Xiaonan Wang wangxiaonan@tsinghua.edu.cn.

Abstract

Materials property prediction is central to high throughput screening, but practical workflows often have few labeled structures and many unlabeled candidates because DFT and experiments are costly. In addition, screening decisions are driven by Top K reliability and ranking stability, not only global MAE or RMSE. We present **SemiMat**, a modular semi supervised regression toolkit tailored to materials property prediction. SemiMat provides a unified pipeline that jointly trains on labeled and unlabeled crystal graphs, supports interchangeable GNN backbones, and exposes configurable components such as structure aware augmentations, confidence based pseudo labels, and teacher student consistency. We implement **RankUp** [1], a ranking aware semi supervised objective that improves Top K robustness by learning pairwise order relations through an auxiliary ranking classifier. Across multiple properties and backbones, SemiMat yields consistent gains over supervised baselines under both regression and screening oriented metrics.

1. Motivation and contribution

High throughput screening pipelines routinely generate large candidate sets, for example from databases or generative proposal models, but only a small fraction can be labeled with DFT or experiments. This makes it hard to train accurate predictors and, more importantly, to produce stable rankings for decision making.

SemiMat is designed for this label scarce setting and emphasizes screening oriented evaluation. Our key contributions are:

- **Unified semi supervised pipeline.** A configurable training and evaluation loop that combines labeled data with an unlabeled structure pool.
- **Ranking aware learning.** RankUp augments regression with an auxiliary ranking classifier (ARC) that learns reliable pairwise order relations and improves Top $K stability$.
- **Practical stability knobs.** We support structure aware weak and strong augmentations, confidence masking for pseudo labels, and an optional EMA teacher for smoother targets.

2. SemiMat toolkit overview

SemiMat aims to make semi supervised learning practical for materials screening. It follows three principles.

Unified labeled and unlabeled training. A single trainer consumes labeled and unlabeled batches in each step, enabling consistent schedules, warm up, and evaluation.

Plug in encoders and learning heads. SemiMat decouples the graph encoder from the learning head. Users can swap encoders such as SchNet [2], MEGNet [3], DimeNet++ [4], equivariant transformer style encoders, and MatGL based encoders [5], while reusing the same semi supervised head and training loop.

Screening oriented metrics. Alongside MAE and NMAE, SemiMat reports screening metrics such as Top K quality and ranking robustness. This supports method selection based on what matters in a screening pipeline.

Figure 1 sketches the overall flow.

3. Related work

Crystal graph models. Graph neural networks are a common choice for periodic crystals because they encode atoms, bonds, and local environments. CGCNN and MEGNet are representative message passing architectures for crystal property prediction [6, 3], and ALIGNN extends the representation with atomistic line graphs [7]. More recently, universal interatomic potentials such as M3GNet and CHGNet aim to learn transferable representations from large scale materials data [8, 9], and MatGL provides an open source stack for training and deploying these models [5]. Large repositories such as the Materials Project and JARVIS enable model development, but high quality labels for target properties remain limited in many practical campaigns [10, 11].

Semi supervised learning. Consistency regularization and pseudo labeling are widely used to exploit unlabeled inputs. Temporal Ensembling and the Π model match predictions under perturbations [12], Mean Teacher uses an EMA teacher to generate smoother targets [13], and MixMatch and FixMatch combine strong augmentation with confidence based pseudo labels [14, 15].

Semi supervised regression and ranking. Semi supervised regression is more sensitive to pseudo label noise because targets are continuous. Recent methods improve stability using uncertainty weighting or representation learning, such as UCVME [16] and CLSS [17]. RankUp converts regression into a ranking problem by introducing an auxiliary ranking classifier, enabling confidence masked consistency training and regression distribution alignment for refining pseudo targets [1].

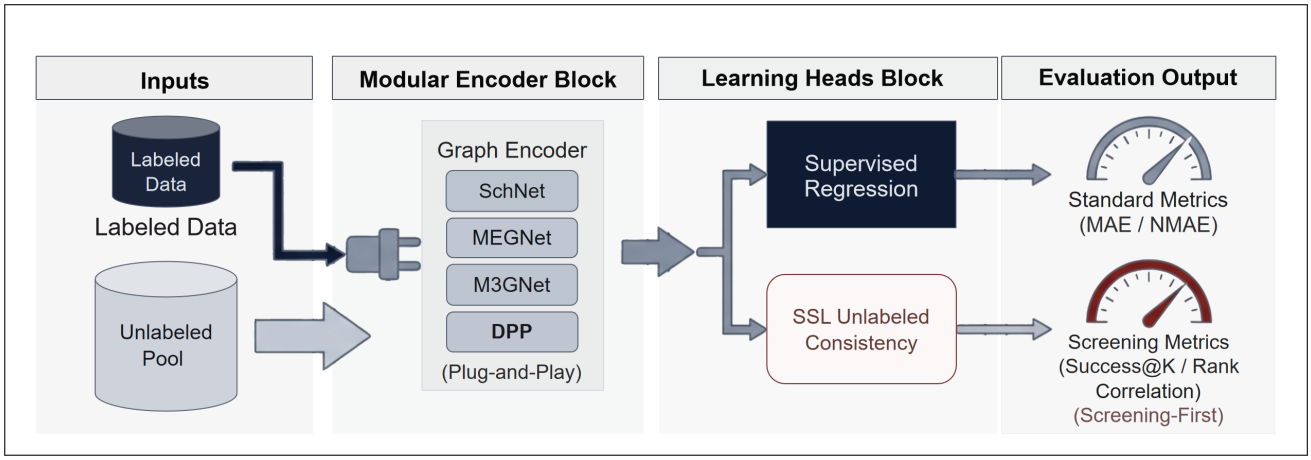

Fig. 1: SemiMat provides a unified pipeline for semi supervised materials property prediction, with modular encoders, structure aware augmentations, confidence based pseudo labels, optional Mean Teacher training, and ranking aware objectives for screening.

4. RankUp and RankUp + Mean Teacher

SemiMat follows an encoder plus pooling plus regression head pattern. RankUp adds an ARC head that predicts pairwise ordering within a mini batch. For labeled pairs, the target is given by the sign of the ground truth difference. For unlabeled structures, RankUp uses weakly augmented views to produce pseudo ordering labels, then enforces consistency on strongly augmented views with confidence masking.

RankUp + Mean Teacher augments RankUp with an EMA teacher network [13]. The teacher provides more stable pseudo targets and supports additional consistency losses, including regression consistency and pairwise consistency. In our implementation, all components are exposed as configuration options, so the same code path can reproduce supervised baselines, standard semi supervised baselines, and ranking aware variants.

5. Evaluation summary

We evaluate on six property prediction tasks including JARVIS based DFT datasets and experimental property datasets [11]. We also use an unlabeled pool (mp-5k) drawn from Materials Project structures [10]. We compare against supervised learning and common semi supervised baselines adapted to regression, including the Π model, Mean Teacher, and MixMatch [12, 13, 14].

Table 1: Overall results averaged over backbones and datasets. Lower Avg NMAE and Avg Rank are better; higher Win Rate is better.

Method	Avg NMAE↓	Avg Rank↓	Win Rate↑
Supervised	0.9284	5.65	0.00
Pi Model	0.9090	3.90	0.05
Mean Teacher	0.8882	2.50	0.25
MixMatch	0.9121	4.15	0.00
RDA	0.9658	4.85	0.20
CLSS	0.9256	4.85	0.10
RankUp (SemiMat)	**0.8845**	**2.10**	**0.40**

Results analysis. Compared with supervised training, RankUp improves screening oriented metrics substantially (Avg Rank from 5.65 to 2.10 and Win Rate from 0.00 to 0.40 in Table 1), indicating more reliable Top K selection across datasets and backbones. The improvement in Avg NMAE is smaller but consistent, suggesting that the auxiliary ranking classifier mainly strengthens relative ordering while preserving regression accuracy. Mean Teacher reduces regression error but is less consistent on ranking summaries, which highlights the benefit of confidence masked ranking consistency for semi supervised screening.

6. Conclusion

SemiMat targets the practical regime of few labeled materials and many unlabeled candidate structures. By combining semi supervised learning with a ranking aware auxiliary objective, RankUp improves both regression accuracy and screening robustness. The toolkit is modular and can be paired with candidate generation to support generate, screen, selectively label, and retrain loops for materials discovery.

References

[1] Pin-Yen Huang, Szu-Wei Fu, and Yu Tsao. Rankup: Boosting semi-supervised regression with an auxiliary ranking classifier. In *Advances in Neural Information Processing Systems*, 2024.

[2] Kristof T. Schütt, Huziel E. Sauceda, Pieter-Jan Kindermans, Alexandre Tkatchenko, and Klaus-Robert Müller. Schnet: A deep learning architecture for molecules and materials. *The Journal of Chemical Physics*, 148(24):241722, 2018.

[3] Chi Chen, Weike Ye, Yunxing Zuo, Chen Zheng, and Shyue Ping Ong. Graph networks as a universal machine learning framework for molecules and crystals. *Chemistry of Materials*, 31(9):3564–3572, 2019.

[4] Johannes Klicpera, Janek Groß, and Stephan Günnemann. Directional message passing for

molecular graphs. In *International Conference on Learning Representations*, 2020.

[5] Tsz Wai Ko, Bowen Deng, Marcel Nassar, et al. Materials graph library (matgl), an open-source graph deep learning library for materials science and chemistry. *npj Computational Materials*, 11:253, 2025.

[6] Tian Xie and Jeffrey C. Grossman. Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties. *Physical Review Letters*, 120(14):145301, 2018.

[7] Kamal Choudhary and Brian DeCost. Atomistic line graph neural network for improved materials property predictions. *npj Computational Materials*, 7:185, 2021.

[8] Chi Chen and Shyue Ping Ong. A universal graph deep learning interatomic potential for the periodic table. *Nature Computational Science*, 2(11):718–728, 2022.

[9] Bowen Deng, Peichen Zhong, KyuJung Jun, et al. Chgnet as a pretrained universal neural network potential for charge-informed atomistic modelling. *Nature Machine Intelligence*, 5:1031–1041, 2023.

[10] Anubhav Jain, Shyue Ping Ong, Geoffroy Hautier, Wei Chen, William Davidson Richards, Stephen Dacek, Shreyas Cholia, Dan Gunter, David Skinner, Gerbrand Ceder, and Kristin A. Persson. Commentary: The materials project: A materials genome approach to accelerating materials innovation. *APL Materials*, 1(1):011002, 2013.

[11] Kamal Choudhary, Kevin F. Garrity, Andrew C. E. Reid, et al. The joint automated repository for various integrated simulations (jarvis) for data-driven materials design. *npj Computational Materials*, 6:173, 2020.

[12] Samuli Laine and Timo Aila. Temporal ensembling for semi-supervised learning. In *International Conference on Learning Representations*, 2017.

[13] Antti Tarvainen and Harri Valpola. Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results. In *Advances in Neural Information Processing Systems*, pages 1195–1204, 2017.

[14] David Berthelot, Nicholas Carlini, Ian Goodfellow, Nicolas Papernot, Avital Oliver, and Colin Raffel. Mixmatch: A holistic approach to semi-supervised learning. In *Advances in Neural Information Processing Systems*, 2019.

[15] Kihyuk Sohn, David Berthelot, Chun-Liang Li, Zizhao Zhang, Nicholas Carlini, Ekin D. Cubuk, Alex Kurakin, Han Zhang, and Colin Raffel. Fixmatch: Simplifying semi-supervised learning with consistency and confidence. In *Advances in Neural Information Processing Systems*, 2020.

[16] Weihang Dai, Xiaomeng Li, and Kwang-Ting Cheng. Semi-supervised deep regression with uncertainty consistency and variational model ensembling via bayesian neural networks. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 2023.

[17] Weihang Dai, Xiaomeng Li, and Kwang-Ting Cheng. Semi-supervised contrastive learning for deep regression. In *Advances in Neural Information Processing Systems*, 2023.