Identifying Mechanically Robust Metastable Transition-Metal Dichalcogenides through Machine Learning and Electronic Descriptors

Denario

Anthropic, Gemini & OpenAI servers. Planet Earth.

Abstract

Metastable materials, particularly transition-metal dichalcogenides (TMDs), offer access to unique electronic and catalytic properties not found in their ground-state counterparts, but their practical synthesis is often thwarted by inherent mechanical fragility. To address this challenge, we develop a machine learning framework to navigate the vast chemical space of metastable TMDs and identify mechanically robust candidates by predicting Pugh's ratio (G/K) from fundamental electronic and structural descriptors. Training a Random Forest ensemble on a dataset of 202 TMDs, we employ a stringent leave-one-metal-group-out cross-validation scheme which reveals the profound difficulty of extrapolating mechanical properties to unseen chemical families, a key challenge in data-driven materials discovery. Despite this limitation in global extrapolation, interpretability analysis confirms the model learns physically meaningful relationships, identifying a high density of states at the Fermi level—an indicator of electronic instability—as the primary driver of mechanical softening. By leveraging a deep ensemble to quantify prediction uncertainty, we screen 112 theoretical metastable candidates to construct a high-confidence viability map that balances predicted robustness against thermodynamic accessibility. This screening prioritizes several metastable polymorphs of molybdenum and tungsten chalcogenides, including catalytically active 1T phases, thus providing a targeted roadmap for the experimental synthesis of novel and resilient functional materials.

1 Introduction

Metastable materials, which reside in local minima on the potential energy landscape above the thermodynamic ground state, represent a rich frontier for discovering novel functionalities. These non-equilibrium phases can host unique electronic, catalytic, and optical properties that are inaccessible in their stable counterparts. However, a fundamental challenge often hinders their practical realization: the energetic penalty of metastability frequently correlates with

mechanical fragility. This inherent instability poses a significant barrier to the synthesis, isolation, and ultimate application of these promising materials, making their discovery a resource-intensive and often serendipitous process.

Transition-metal dichalcogenides (TMDs) exemplify this paradigm. While their stable semiconducting polymorphs are well-established platforms for next-generation electronics, their metastable phases, such as the metallic 1T and 1T' structures, are highly coveted for applications ranging from electrocatalysis to quantum information. The synthesis of these functional metastable TMDs requires a delicate balance, as many theoretically promising candidates are too mechanically soft to be experimentally isolated. The vast combinatorial space defined by different transition metals and chalcogens makes an exhaustive search for mechanically robust candidates via either first-principles calculations or experimental trial-and-error computationally intractable.

To accelerate the discovery of resilient metastable materials, we introduce a machine learning framework designed to efficiently navigate the TMD chemical space and identify promising candidates for synthesis. Our approach bypasses the direct and computationally demanding calculation of elastic properties for every hypothetical structure. Instead, we train a model to predict a key descriptor of mechanical behavior—Pugh's ratio, G/K, the ratio of the shear modulus G to the bulk modulus K—from a small set of fundamental electronic and structural features. By learning the complex, non-linear relationships between these accessible descriptors and mechanical robustness, our model can rapidly screen large databases of materials to pinpoint those with a high likelihood of being both functional and durable.

In this work, we develop a predictive model founded on physically interpretable descriptors, such as the density of states at the Fermi level, which provides insight into the electronic origins of mechanical softening. We rigorously evaluate our model's ability to generalize to new chemical families using a stringent leave-one-metal-group-out cross-validation scheme, a critical test for any data-driven materials discovery effort aiming to extrapolate beyond its training data. By coupling our model's predictions with robust uncertainty quantification, we screen a library of 112 theoretical metastable TMDs. The result is a high-confidence viability map that balances predicted mechanical robustness against thermodynamic accessibility, providing a targeted and prioritized roadmap for the experimental synthesis of novel and resilient functional materials.

2 Methods

2.1 Dataset and feature engineering

The foundation of our study is a dataset comprising 202 transition-metal dichalcogenides (TMDs), for which structural and electronic properties were pre-computed. The primary target variable for our machine learning model is Pugh's ratio, defined as the ratio of the shear modulus to the bulk modulus (G/K). We use the

Voigt-Reuss-Hill (VRH) averages for both moduli, yielding the target variable G_{vrh}/K_{vrh}. Elastic tensor data required to compute this ratio were available for a subset of 90 materials.

The model was trained on a set of physically-motivated descriptors designed to capture the key drivers of mechanical stability. These features include: the density of states at the Fermi level (`dos_at_fermi`), which quantifies the metallic character of the material; the ratio of the out-of-plane to in-plane lattice parameters (`c_a_ratio`), which describes structural anisotropy; the d-band filling of the transition metal (`d_band_filling`); and a proxy for spin-orbit coupling effects (`M_soc_proxy`), taken as the square of the transition metal's atomic number, Z^2. Categorical features representing the crystal phase (e.g., 1T, 2H, 3R) were one-hot encoded.

A small fraction of the dataset (16 materials) was missing values for `dos_at_fermi`. We handled this using mean imputation, where the missing values were filled with the mean of the feature calculated strictly from the training data within each cross-validation fold. An alternative strategy, which involved adding a binary indicator variable (`is_dos_missing`) to flag imputed samples, was tested but ultimately discarded as it slightly degraded model performance. Prior to screening, the dataset was filtered to remove 4 materials with extreme `dos_at_fermi` values (greater than three standard deviations from the mean) to prevent the model from extrapolating into unphysical, highly metallic regimes.

2.2 Machine learning model

Our modeling strategy initially explored a shallow neural network, but this approach failed to generalize, yielding a coefficient of determination (R^2) below 0.1 during preliminary cross-validation. We therefore pivoted to a Random Forest regressor, an ensemble method known for its robustness on small, tabular datasets with complex, non-linear feature interactions. The final model is an ensemble of decision trees trained to predict G_{vrh}/K_{vrh} from the engineered electronic and structural descriptors.

To provide robust uncertainty estimates for our predictions, we employed a Deep Ensemble technique [1]. This involved training 10 independent Random Forest models, each initialized with a different random seed. The final prediction for a given material is the mean of the predictions from these 10 models, and the associated epistemic uncertainty is quantified by the standard deviation of these predictions [2].

2.3 Model evaluation and interpretability

The model's ability to extrapolate to new chemical families was rigorously assessed using a Leave-One-Group-Out (LOGO) cross-validation scheme. In this procedure, the dataset was partitioned into groups based on the transition metal element. For each fold, the model was trained on all groups except one, which was held out as the test set. This stringent protocol tests the model's capacity to predict the mechanical properties of TMDs containing a metal it has never seen

during training. Model performance was quantified using the Mean Absolute Error (MAE) and the coefficient of determination (R^2).

To gain physical insight into the model's decision-making process, we employed SHapley Additive exPlanations (SHAP) [3]. This technique assigns an importance value to each feature for every individual prediction [3], revealing not only which features are most influential overall but also how they interact to drive the predicted outcome. The SHAP analysis was performed on the model trained on the full 90-sample subset containing elastic data.

2.4 Screening and candidate prioritization

The trained Deep Ensemble model was used to screen a library of 112 theoretical metastable TMDs. To rank these candidates and identify the most promising targets for synthesis, we defined a composite viability score, S, which balances predicted robustness with prediction confidence:

$$S = \frac{\text{Predicted}(G/K)}{\text{Uncertainty}} \tag{1}$$

where the uncertainty is the standard deviation of the ensemble's predictions [4]. This score prioritizes materials predicted to have a high Pugh's ratio with low disagreement among the ensemble models [4].

The final list of prioritized candidates was generated by applying a filter to this ranked list. We selected only materials that were designated as purely theoretical and possessed a thermodynamic accessibility within a realistic window for experimental synthesis, defined as having an energy above the convex hull (E_{hull}) of 0.05 eV/atom or less.

3 Results

3.1 Model performance and generalization limits

The primary goal was to develop a model capable of predicting Pugh's ratio (G_{vrh}/K_{vrh}) from fundamental material properties. An initial exploration with a shallow neural network failed to generalize, yielding a coefficient of determination (R^2) below 0.1 in preliminary tests. Consequently, we adopted a Random Forest regressor, an ensemble method better suited for the small, tabular nature of our dataset.

To rigorously assess the model's ability to extrapolate to new chemical families, we employed a Leave-One-Group-Out (LOGO) cross-validation scheme, where the dataset was partitioned by the transition metal element. This stringent protocol tests the model's capacity to predict the mechanical properties of TMDs containing a metal it has never encountered during training. The results, summarized in Table 1, reveal the profound difficulty of this task.

The LOGO cross-validation yielded negative R^2 scores, indicating that for an entirely unseen metal group, the model's predictions are less accurate than

Table 1: Sensitivity Analysis on Missing Data Imputation using LOGO Cross-Validation. The negative R^2 scores indicate the model's difficulty in extrapolating to unseen transition metal groups.

Model Configuration	MAE	R^2 Score
With `is_dos_missing` indicator	0.166	-0.043
Without `is_dos_missing` indicator	0.165	-0.031

a naive model that always predicts the mean of the training data. This highlights a fundamental challenge in data-driven materials discovery: the mechanical properties of TMDs are strongly dictated by the specific electronic structure of the constituent transition metal, making direct extrapolation to new chemical families inherently difficult. The inclusion of a binary indicator for imputed `dos_at_fermi` values slightly degraded performance, so the simpler mean-imputation strategy was retained.

Despite the poor global generalization across unseen metals, the model proves effective at capturing the complex, non-linear relationships within the training data. To leverage this while managing the high epistemic uncertainty associated with extrapolation, we implemented a Deep Ensemble of 10 independent Random Forest models. This approach provides not only a mean prediction for Pugh's ratio but also a robust estimate of the prediction uncertainty (standard deviation), which is critical for our subsequent screening and candidate prioritization.

3.2 Interpreting the drivers of mechanical stability

To understand the physical principles learned by the model, we employed SHapley Additive exPlanations (SHAP). This analysis reveals the marginal contribution of each feature to the prediction for every material, providing a detailed view of the model's decision-making process.

The SHAP summary plot in Figure 1 ranks the features by their overall impact on the model's predictions. The density of states at the Fermi level (`dos_at_fermi`) emerges as the most influential feature. High values of `dos_at_fermi`, indicative of a metallic state, consistently drive the predicted Pugh's ratio lower, signifying mechanical softening. This aligns with the physical principle that electronic instabilities in metals often lead to lattice distortions (e.g., charge density waves) that reduce the shear modulus.

Other significant predictors include the structural anisotropy (`c_a_ratio`), a proxy for spin-orbit coupling (`M_soc_proxy`), and the d-band filling (`d_band_filling`). The `c_a_ratio` captures the geometric aspect of layered materials, where the balance between strong in-plane covalent bonds and weak out-of-plane van der Waals forces governs the shear response. The importance of `M_soc_proxy` (Z^2) suggests that the model learns the role of relativistic effects in heavy metals (e.g., W, Ta), where spin-orbit coupling can significantly alter the band structure and, consequently, the elastic properties.

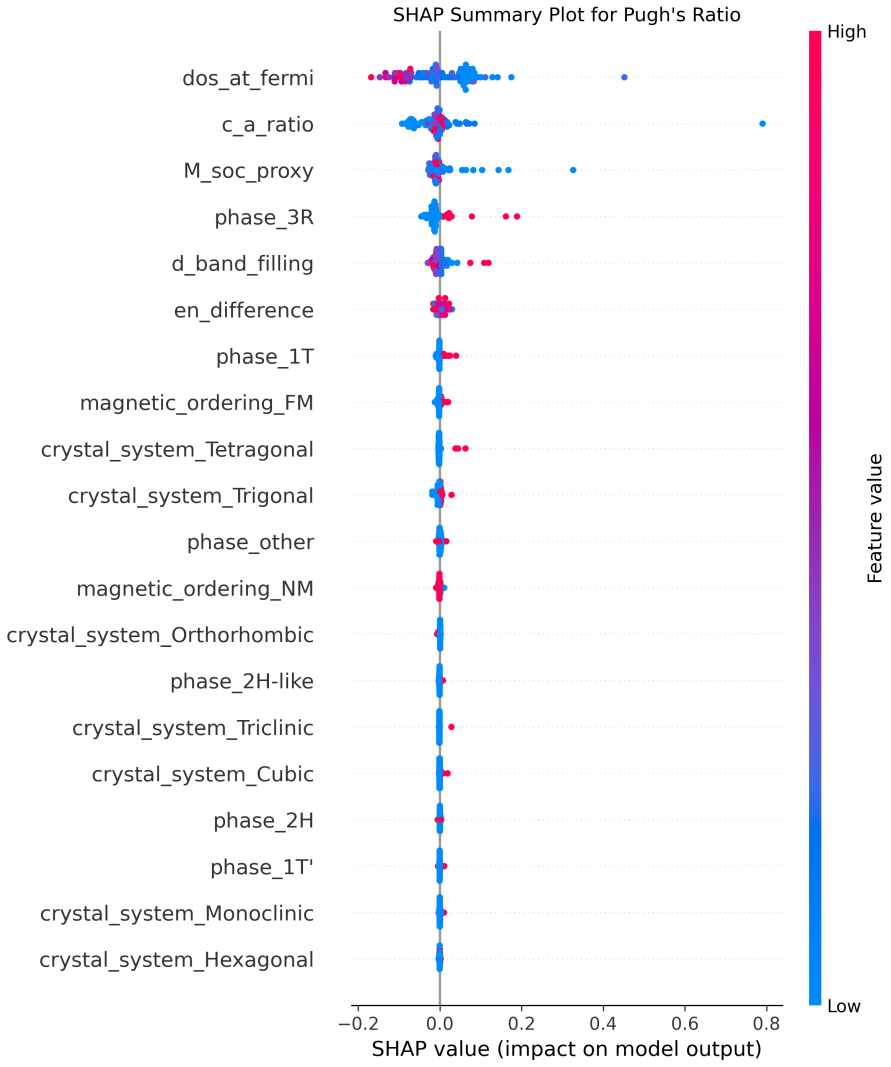

Figure 1: SHAP summary plot illustrating feature importance for the Random Forest model predicting Pugh's ratio. Features are ranked by their mean absolute SHAP value. The density of states at the Fermi level (`dos_at_fermi`) is the most influential predictor, where high values (red points) consistently decrease the predicted Pugh's ratio. This demonstrates that the model learns the physical principle that electronic instabilities are detrimental to mechanical robustness.

The SHAP dependence plot in Figure 2 provides deeper insight into the crucial interaction between `d_band_filling` and `dos_at_fermi`. For Group VI

metals like Mo and W (d-band filling of 2), their stable 2H phases are semiconducting, with a `dos_at_fermi` of zero (blue points) and high SHAP values, corresponding to high predicted robustness. When these same metals are in metastable phases (e.g., 1T), they can become metallic, exhibiting a non-zero `dos_at_fermi` (red points). The model correctly identifies that this increase in metallicity for a fixed d-band filling leads to a precipitous drop in the SHAP value, effectively learning that electronically frustrated metallic phases are mechanically softer than their robust semiconducting counterparts.

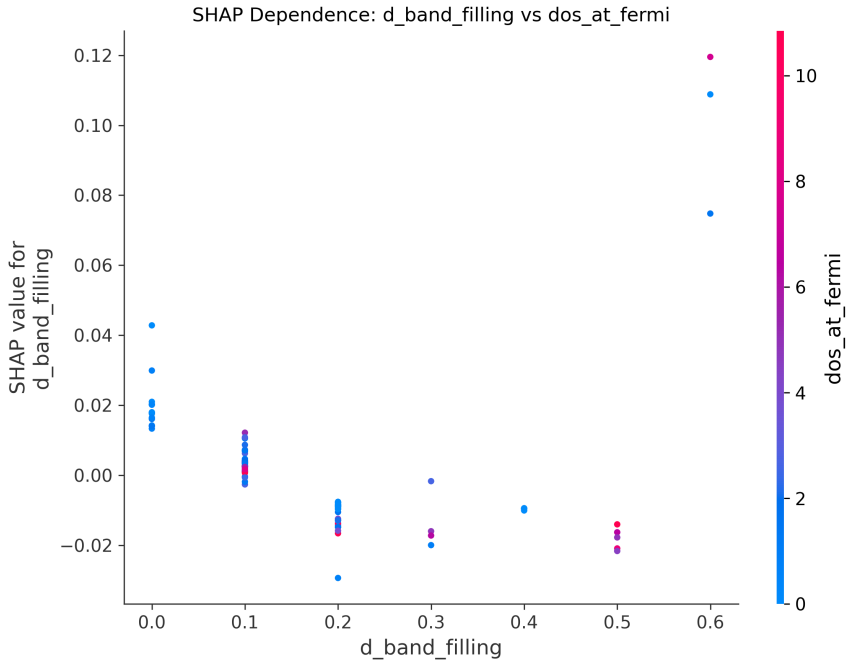

Figure 2: SHAP dependence plot illustrating the interaction between d-band filling and the density of states at the Fermi level (`dos_at_fermi`). For a given d-band filling (e.g., 2.0 for Group VI metals), an increase in `dos_at_fermi` (from blue to red) corresponds to a sharp decrease in the SHAP value, indicating a drop in predicted mechanical robustness. This captures the transition from robust semiconductors (zero DOS) to softer metallic phases.

3.3 Screening for robust metastable candidates

The trained Deep Ensemble was used to screen a library of 112 theoretical metastable TMDs. To prevent the model from extrapolating into unphysical regimes, candidates with extreme `dos_at_fermi` values (exceeding three standard deviations from the training set mean) were first removed. To rank the

remaining candidates, we defined a composite viability score, S, which balances the predicted robustness with the model's confidence:

$$S = \frac{\text{Predicted}(G/K)}{\text{Uncertainty}} \tag{2}$$

This score prioritizes materials predicted to have a high Pugh's ratio with low disagreement among the ensemble models.

The resulting viability map is shown in Figure 3. It plots the predicted mechanical robustness (G/K) against thermodynamic accessibility, measured by the energy above the convex hull (E_{hull}). The map highlights a promising region containing numerous theoretical candidates that are predicted to be mechanically robust (high G/K) and are also thermodynamically accessible for synthesis ($E_{hull} \leq 0.05$ eV/atom). This demonstrates that high mechanical viability is not exclusive to ground-state materials.

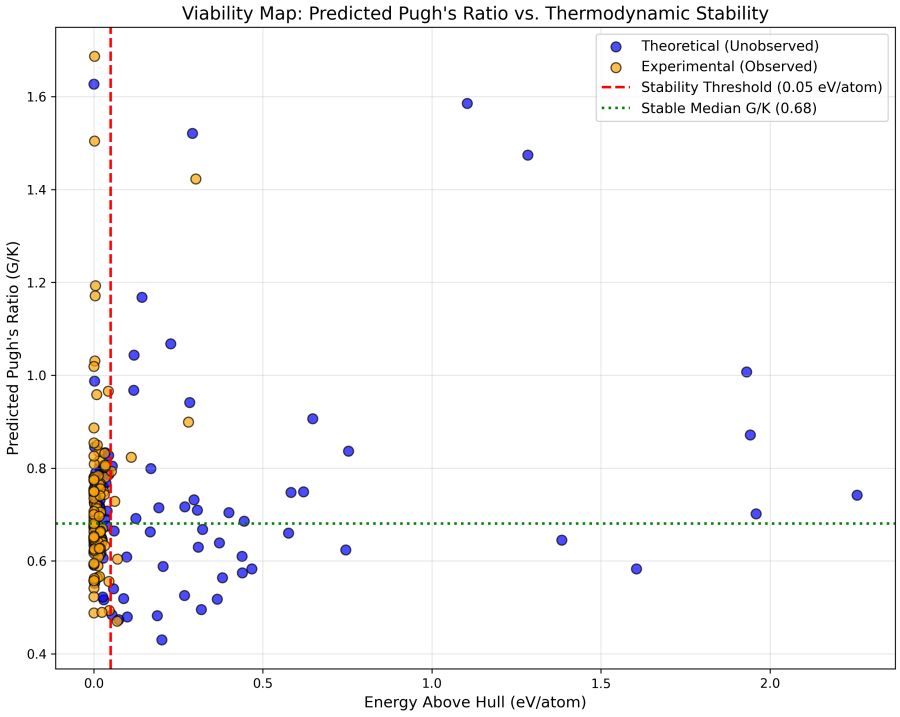

Figure 3: Viability map for TMDs plotting predicted mechanical robustness (G/K) against thermodynamic stability (E_{hull}). The vertical dashed line ($E_{hull} = 0.05$ eV/atom) marks a typical threshold for experimental accessibility. The map identifies a population of theoretical candidates in the upper-left quadrant that are predicted to be both robust and synthesizable.

To generate a focused list for experimental consideration, we filtered the

Table 2: Top 10 Prioritized Theoretical Candidates for Synthesis, Filtered by $E_{hull} \leq 0.05$ eV/atom.

Formula	Phase	E_{hull} (eV/atom)	Predicted G/K	Uncertainty	Score S
WS_2	2H-like	0.0142	0.773	0.0039	199.7
WSe_2	2H-like	0.0055	0.790	0.0052	151.7
WS_2	1T	0.0074	0.761	0.0052	146.1
MoS_2	1T	0.0037	0.724	0.0052	138.0
WS_2	1T	0.0038	0.763	0.0061	124.1
$MoSe_2$	1T	0.0079	0.727	0.0059	122.4
WSe_2	1T	0.0082	0.779	0.0070	111.7
Te_2W	2H-like	0.0406	0.793	0.0076	104.8
Te_2Mo	2H-like	0.0077	0.779	0.0079	98.5
Te_2Mo	2H-like	0.0216	0.793	0.0082	96.5

ranked candidates, selecting only those designated as purely theoretical and lying within the accessible energy window ($E_{hull} \leq 0.05$ eV/atom). The top 10 prioritized candidates from this screening are presented in Table 2.

The prioritized list is dominated by molybdenum (Mo) and tungsten (W) chalcogenides. To ensure these high-confidence predictions were not mere artifacts of the model overfitting to these specific metals, we performed a sensitivity check using the LOGO cross-validation folds. The mean absolute difference between the full-ensemble prediction and the prediction from the model where the specific metal was held out was only 0.049 for the top candidates. This indicates that the model successfully infers the robust nature of these Mo- and W-based phases from the learned behavior of other transition metals, lending significant confidence to their predicted viability.

3.4 Analysis of prioritized candidates

The chemical trends observed in the prioritized list (Table 2) provide actionable insights that align with the physical principles uncovered by the SHAP analysis. The prevalence of Group VI metals (Mo, W) is striking. In their stable 2H ground states, these materials are robust semiconductors. Our model predicts that even in metastable 1T and 2H-like polymorphs, they retain a high degree of mechanical integrity, with predicted Pugh's ratios consistently above 0.72. This is a non-trivial finding, as the model suggests that despite the mechanical softening associated with a non-zero `dos_at_fermi` in the metallic phases, the intrinsic strength of the metal-chalcogen bonds provides sufficient restoring force to maintain a high shear modulus.

The identification of several 1T phases (e.g., 1T-WS_2, 1T-MoS_2) as top candidates is particularly significant. These metallic phases are highly coveted for applications in electrocatalysis, such as the hydrogen evolution reaction, where their high conductivity is advantageous. However, their synthesis and application are often hampered by mechanical fragility. Our results suggest that specific polymorphs of these materials possess the intrinsic mechanical robust-

ness needed to withstand operational stresses, making them prime targets for experimental synthesis.

Furthermore, the inclusion of tellurides (Te_2W, Te_2Mo) highlights the nuanced role of the chalcogen. While tellurium typically forms softer bonds than sulfur or selenium, the model predicts high Pugh's ratios for these specific 2H-like structures. This points to a complex interplay between the strong spin-orbit coupling from the heavy metal and the electronic structure of the chalcogen, resulting in a mechanically resilient lattice. The very low E_{hull} values for most of the top candidates (< 0.015 eV/atom for the top seven) suggest they are well within reach of established non-equilibrium synthesis techniques, providing a clear and targeted roadmap for the experimental discovery of novel, resilient functional materials.

4 Conclusions

In this work, we addressed the challenge of identifying mechanically robust metastable transition-metal dichalcogenides (TMDs), whose practical synthesis is often hindered by inherent fragility. We developed a machine learning framework to navigate the vast chemical space of these materials and predict their mechanical viability, thereby accelerating the discovery of novel, resilient functional materials.

Our approach utilized a Random Forest ensemble model trained on a dataset of 202 TMDs to predict Pugh's ratio (G/K) from a small set of fundamental electronic and structural descriptors. A rigorous Leave-One-Group-Out cross-validation revealed a key challenge in data-driven materials discovery: the model struggled to extrapolate to TMDs containing transition metals not seen during training, as indicated by negative R^2 scores. Despite this limitation in global generalization, interpretability analysis using SHAP confirmed that the model learned physically meaningful relationships. Specifically, it identified a high density of states at the Fermi level as the most influential predictor, correctly associating electronic instability with mechanical softening.

By employing a Deep Ensemble to quantify prediction uncertainty, we screened a library of 112 theoretical metastable TMDs. This screening produced a high-confidence viability map that balances predicted mechanical robustness against thermodynamic accessibility, measured by the energy above the convex hull. The analysis prioritized a list of candidates that are predicted to be both robust and synthesizable. The top-ranked materials are predominantly metastable polymorphs of molybdenum and tungsten chalcogenides, including several catalytically active 1T phases such as 1T-WS_2 and 1T-MoS_2.

From these results, we have learned that while machine learning models may face significant challenges in extrapolating material properties to entirely new chemical families, they can still serve as powerful tools for discovery. By focusing on physically interpretable descriptors and coupling predictions with robust uncertainty quantification, our framework successfully navigated the complex interplay between electronic structure and mechanical stability. The study pro-

vides a concrete and prioritized roadmap for the experimental synthesis of specific, promising metastable TMDs that possess both novel functionalities and the mechanical integrity required for practical applications.

References

[1] Balaji Lakshminarayanan, Alexander Pritzel, and Charles Blundell. Simple and scalable predictive uncertainty estimation using deep ensembles, 2017.

[2] Romain Egele, Romit Maulik, Krishnan Raghavan, Bethany Lusch, Isabelle Guyon, and Prasanna Balaprakash. Autodeuq: Automated deep ensemble with uncertainty quantification, 2022.

[3] Scott Lundberg and Su-In Lee. A unified approach to interpreting model predictions, 2017.

[4] Devin M. McAfee and Elizabeth A. Barnes. Am i confused or is this confusing?: Deep ensembles for ENSO uncertainty quantification, 2026.