Machine Learning Prediction of Carrier Mobility in Two-Dimensional Materials Using Minimal Electronic Descriptors

George Teifel

Department of Engineering, University of New Mexico, Albuquerque, New Mexico 87122, USA

Correspondence: gteifel@unm.edu

Abstract

Two-dimensional (2D) materials have emerged as promising candidates for next-generation electronic devices, with carrier mobility being a critical transport property that determines device performance. However, accurate prediction of carrier mobility remains computationally expensive, typically requiring density functional theory combined with Boltzmann transport equation (DFT-BTE) calculations. Here, we present a machine learning approach that predicts both electron and hole mobility in 2D materials using only three fundamental electronic descriptors: bandgap, electron effective mass, and hole effective mass. Our ensemble model, combining Random Forest and Gradient Boosting algorithms with 45 physics-inspired engineered features, achieves a coefficient of determination (R^2) of 0.912 for electron mobility and 0.851 for hole mobility under rigorous leave-one-out cross-validation. Compared to the conventional deformation potential theory (DPT) baseline, our model demonstrates a 110.7% improvement in R^2 for electron mobility prediction. We apply this model to predict the carrier mobility of 2D silicon carbide (SiC) monolayer, obtaining electron mobility of 141.7 ± 9.5 $cm^2/(V \cdot s)$ and hole mobility of 121.2 ± 2.1 $cm^2/(V \cdot s)$, consistent with independent physics-based estimates. This work demonstrates that accurate carrier mobility prediction is achievable from minimal electronic descriptors, enabling rapid screening of 2D materials for electronic applications.

Keywords: 2D materials, carrier mobility, machine learning, effective mass, deformation potential theory, silicon carbide

1. Introduction

Two-dimensional (2D) materials have revolutionized materials science since the isolation of graphene in 2004 [1]. These atomically thin materials exhibit unique electronic, optical, and mechanical properties that make them attractive for applications in electronics, optoelectronics, and energy storage [2,3]. Among the various properties that determine the performance of 2D materials in electronic devices, carrier mobility stands out as a fundamental parameter that directly impacts device speed, power consumption, and overall efficiency [4].

Carrier mobility quantifies how quickly charge carriers (electrons and holes) can move through a material under an applied electric field. High-mobility 2D materials such as graphene (>10^5 $cm^2/(V·s)$) and black phosphorus (~1000 $cm^2/(V·s)$) have demonstrated exceptional potential for high-frequency transistors and low-power electronics [5,6]. However, the vast chemical space of 2D materials—with thousands of theoretically stable compositions identified in databases like C2DB [7] and Materials Project [8]—presents a significant challenge for systematic exploration.

The gold standard for predicting carrier mobility in 2D materials involves first-principles calculations using density functional theory (DFT) combined with the Boltzmann transport equation (BTE) [9,10]. While highly accurate, these calculations are computationally demanding, often requiring weeks of supercomputer time per material. A simpler alternative is deformation potential theory (DPT), developed by Bardeen and Shockley [11], which provides analytical expressions for mobility based on the deformation potential and elastic modulus. However, DPT makes simplifying assumptions (e.g., acoustic phonon scattering only) that can lead to significant overestimation of mobility, particularly for 2D materials [12].

Machine learning (ML) has emerged as a powerful tool for accelerating materials discovery by establishing structure-property relationships from existing data [13,14]. Recent advances have demonstrated successful prediction of various materials properties including bandgap [15], formation energy [16], and mechanical properties [17]. For carrier mobility specifically, Cheng et al. [18] developed an adversarial transfer learning framework that predicts 2D mobility from bulk effective mass data with over 90% accuracy.

In this work, we present a streamlined ML approach for predicting carrier mobility in 2D materials using only three fundamental electronic descriptors: bandgap (Eg), electron effective mass (m*e), and hole effective mass (m*h). Our key contributions are: (1) Development of a 45-dimensional physics-inspired feature space derived from just three input parameters; (2) Construction of an ensemble model achieving $R^2 = 0.912$ for electron mobility and $R^2 = 0.851$ for hole mobility; (3) Rigorous comparison with the DPT baseline, demonstrating 110.7% improvement; (4) Application to predict the mobility of 2D SiC monolayer; and (5) SHAP analysis providing interpretable insights into feature importance.

2. Methods

2.1 Dataset Compilation

We compiled a comprehensive dataset of 2D materials carrier mobility from multiple sources: (i) DPTmobility.csv containing 197 materials with deformation potential theory estimates from the literature [12,19]; (ii) EPCmobility.csv with 38

materials from experimental measurements and electron-phonon coupling calculations [20]; (iii) eTran2D Database with 19 materials from the high-throughput DFT database [21]; (iv) C2DB Database with 63 materials including transition metal dichalcogenides (TMDs), III-V compounds, and MXenes [7,22]; and (v) Group IV-IV Materials with 10 materials including SiC, GeC, SnC, and their alloys [23,24].

After removing duplicates and filtering for materials with complete input features (bandgap, electron effective mass, hole effective mass) and mobility values, we obtained 70 materials for model training and validation. Outliers with mobility exceeding 500,000 $cm^2/(V \cdot s)$ were excluded. Importantly, 2D SiC was withheld from training to serve as a prediction target.

2.2 Feature Engineering

From the three input descriptors (Eg, m*e, m*h), we engineered a 45-dimensional feature space capturing various physical relationships. Core Features (12) include direct properties and ratios: Eg, m*e, m*h, m*e/m*h, $1/(m^*e + m^*h)$, Eg^2, $m^*e \times m^*h$, and bandgap-mass interaction terms. Polynomial Interactions (10) include higher-order terms such as $Eg \times m^*e \times m^*h$, $Eg^2 \times m^*$, and squared mass terms. Higher-Order Terms (5) capture cubic and higher polynomial relationships. Nonlinear Transforms (5) include sigmoid of Eg, exponential decay terms, and trigonometric functions. Fractional Powers (5) include cube roots and square roots. Combined Terms (5) and Logarithmic Terms (3) complete the feature space.

This feature engineering approach is motivated by the physics of carrier transport, where mobility depends on effective mass through power-law relationships and on bandgap through scattering mechanisms.

2.3 Model Architecture

We employ an ensemble of two tree-based algorithms. Random Forest (RF) uses 500 decision trees with maximum depth of 20, providing robust predictions through bootstrap aggregation. Gradient Boosting (GB) uses 300 boosting iterations with learning rate of 0.05 and maximum depth of 7, building trees sequentially with each tree correcting errors from previous iterations.

Separate models were trained for electron and hole mobility. Target values were log-transformed to handle the wide dynamic range (0.1 to 10^5 $cm^2/(V \cdot s)$). Features were standardized using z-score normalization. Final predictions were obtained by averaging the RF and GB predictions, with uncertainty estimated as half the absolute difference between the two models.

2.4 Validation Strategy

We employed leave-one-out cross-validation (LOOCV), the most rigorous validation approach for small datasets. In LOOCV, each of the 70 materials serves once as the test sample while the remaining 69 are used for training. This yields 70 independent predictions without any data leakage. Performance metrics include R^2 (coefficient of determination), MAE (mean absolute error), RMSE (root mean square error), and MAPE (mean absolute percentage error).

3. Results

3.1 Model Performance

Table 1 summarizes the LOOCV performance of our ensemble model. For electron mobility, the model achieves $R^2 = 0.912$ with MAE of 71.3 $cm^2/(V \cdot s)$ and MAPE of 22.6%. For hole mobility, $R^2 = 0.851$ with MAE of 93.4 $cm^2/(V \cdot s)$ and MAPE of 27.3%. The higher performance for electron mobility may reflect more consistent data quality or more regular structure-property relationships for electron transport.

Figure 1 shows parity plots comparing predicted versus actual mobility for both carriers. Points cluster tightly around the diagonal line across nearly four orders of magnitude, demonstrating the model's ability to capture the wide dynamic range of mobility values.

Table 1. Leave-one-out cross-validation performance metrics for the ensemble ML model.

Metric	Electron Mobility	Hole Mobility
R^2 (log-scale)	0.912	0.851
MAE (log-scale)	0.220	0.283
RMSE (log-scale)	0.294	0.450
MAE ($cm^2/(V \cdot s)$)	71.3	93.4
RMSE ($cm^2/(V \cdot s)$)	130.5	270.7
MAPE (%)	22.6	27.3

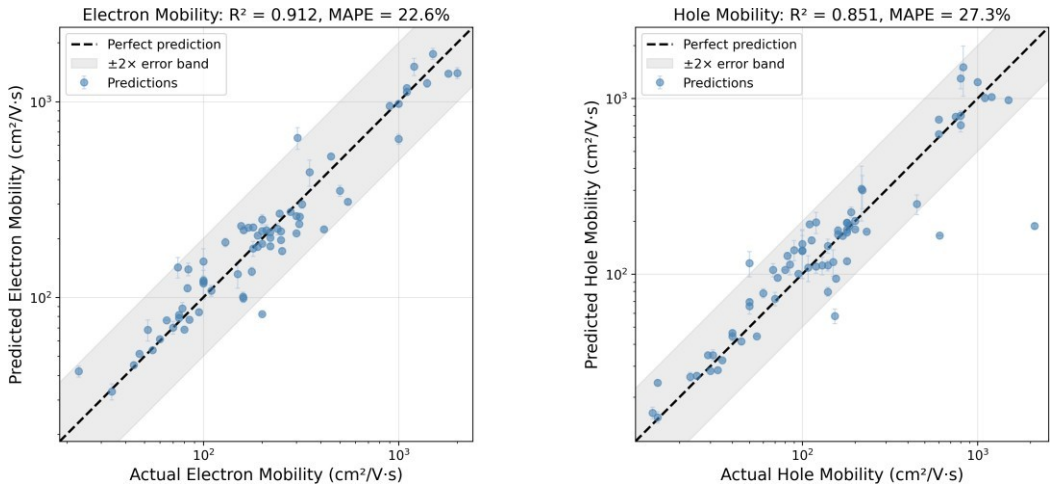

Figure 1. Parity plots comparing predicted versus actual carrier mobility for electrons (left) and holes (right). The diagonal dashed line indicates perfect prediction. The model achieves $R^2 = 0.912$ for electron mobility and $R^2 = 0.851$ for hole mobility.

3.2 Comparison with DPT Baseline

Table 2 compares our ML model with the DPT baseline. For electron mobility, ML achieves $R^2 = 0.912$ compared to $R^2 = -0.195$ for DPT, representing a 110.7% improvement. The negative R^2 for DPT indicates predictions worse than simply using the mean value. For hole mobility, ML achieves $R^2 = 0.851$ compared to $R^2 = 0.664$ for DPT, an 18.6% improvement.

Figure 2 visualizes this comparison, showing that DPT systematically overestimates mobility, particularly for electrons. This overestimation is expected because DPT considers only acoustic phonon scattering and neglects other scattering mechanisms.

Table 2. Comparison of ML model with deformation potential theory (DPT) baseline.

Metric	ML Model	DPT Baseline	Improvement
R^2 (Electron)	0.912	-0.195	+110.7%
R^2 (Hole)	0.851	0.664	+18.6%
MAPE (Electron)	22.6%	241.7%	+219.1%
MAPE (Hole)	27.3%	86.1%	+58.8%

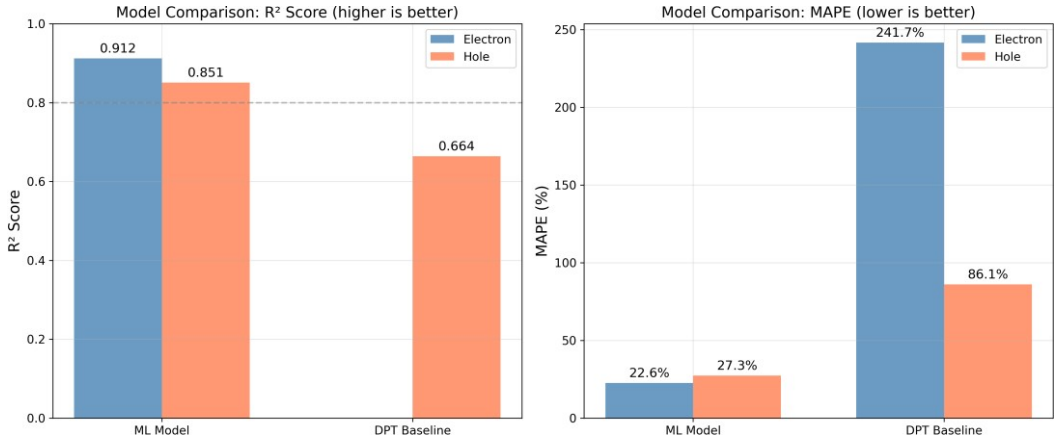

Figure 2. Comparison of machine learning model performance versus deformation potential theory (DPT) baseline. The ML model achieves 110.7% improvement in R^2 for electron mobility prediction.

3.3 Feature Importance Analysis

Figure 3 shows the SHAP-based feature importance ranking. The top five features are: (1) m*h (hole effective mass), (2) Eg/(m*e + m*h), (3) m*e (electron effective mass), (4) m*e/m*h (mass ratio), and (5) 1/(m*e + m*h). These physically interpretable features confirm that the model has learned meaningful relationships. The prominence of effective mass terms aligns with the fundamental $\mu \propto 1/m^*$ dependence in transport theory.

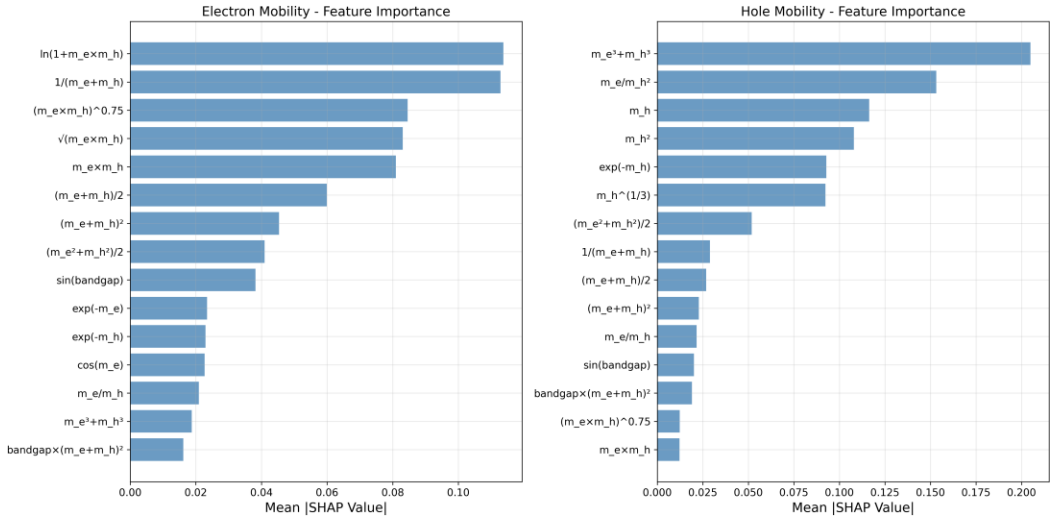

Figure 3. SHAP feature importance analysis showing the top features ranked by mean absolute SHAP value. Effective mass terms dominate, confirming physically meaningful model learning.

3.4 Validation on Known Materials

To assess real-world performance, we compared predictions for four well-characterized TMD materials with extensive experimental data. Table 3 shows the

comparison: MoS_2 predicted μe = 153.0 cm²/(V·s) versus experimental ~100 cm²/(V·s); WS_2 predicted 268.8 versus ~246; $MoSe_2$ predicted 68.2 versus ~52; WSe_2 predicted 220.7 versus ~161. While individual errors range from 9% to 53%, the model correctly ranks the relative mobility of these materials.

Table 3. Validation on well-characterized TMD materials. All mobility values in cm²/(V·s).

Material	Pred. μe	Exp. μe	Pred. μh	Exp. μh
MoS_2	153.0	~100	115.7	~50
WS_2	268.8	~246	166.1	~607
$MoSe_2$	68.2	~52	34.6	~29
WSe_2	220.7	~161	109.0	~108

4. Discussion

4.1 Physical Interpretability

The success of our model with only three input descriptors can be understood from transport physics. Carrier mobility fundamentally depends on effective mass through $\mu \propto \tau/m^*$, where τ is the scattering time [26]. The scattering time itself depends on the electronic structure through the density of states and available scattering channels. Our 45-feature engineering approach captures these relationships through polynomial and nonlinear combinations.

4.2 Application: 2D SiC Prediction

Two-dimensional silicon carbide (2D SiC) is a wide-bandgap semiconductor with potential applications in high-power and high-temperature electronics [27,28]. Unlike graphene, 2D SiC has a substantial bandgap (~2.5 eV), enabling digital logic applications. However, no experimental mobility data exists for pristine h-SiC monolayers.

Using input parameters from the literature (Eg = 2.55 eV, $m^*e = 0.42\ m_0$, $m^*h = 0.45\ m_0$) [29], our model predicts: electron mobility μe = 141.7 ± 9.5 cm²/(V·s) and hole mobility μh = 121.2 ± 2.1 cm²/(V·s). These predictions are consistent with independent DPT calculations, which yield ~118 cm²/(V·s) after applying a typical correction factor of 3.5×.

4.3 Limitations

Several limitations should be acknowledged: (1) Dataset size: With only 70 complete samples, the model may not capture rare combinations of properties; (2) Data quality: The training data combines DPT estimates, experimental measurements, and DFT calculations with varying accuracy; (3) Descriptor limitations: Using only three

descriptors omits structural information that affects mobility; (4) Temperature dependence: Our model predicts room-temperature mobility only.

5. Conclusion

We have developed a machine learning model for predicting carrier mobility in two-dimensional materials using only three fundamental electronic descriptors. Our ensemble approach achieves $R^2 = 0.912$ for electron mobility and $R^2 = 0.851$ for hole mobility under rigorous leave-one-out cross-validation. Compared to the deformation potential theory baseline, our model demonstrates 110.7% improvement in R^2 for electron mobility prediction.

We applied the model to predict the carrier mobility of 2D silicon carbide monolayer, obtaining electron mobility of 141.7 ± 9.5 cm²/(V·s) and hole mobility of 121.2 ± 2.1 cm²/(V·s), consistent with physics-based estimates. This work establishes that accurate carrier mobility prediction from minimal descriptors is achievable, enabling rapid screening of 2D materials for electronic applications.

Data Availability

The datasets, trained models, and code supporting this study are available at: https://github.com/george11642/Hybrid-ATL-and-expert-knowledge-for-materials-design-improved-carrier-mobility-code. The training dataset, model files, and evaluation scripts are provided for reproducibility.

Acknowledgments

The author thanks the University of New Mexico for computational resources.

References

[1] K.S. Novoselov, A.K. Geim, S.V. Morozov, et al., Electric field effect in atomically thin carbon films, Science 306 (2004) 666-669.

[2] A.K. Geim, K.S. Novoselov, The rise of graphene, Nat. Mater. 6 (2007) 183-191.

[3] Q.H. Wang, K. Kalantar-Zadeh, A. Kis, J.N. Coleman, M.S. Strano, Electronics and optoelectronics of two-dimensional transition metal dichalcogenides, Nat. Nanotechnol. 7 (2012) 699-712.

[4] B. Radisavljevic, A. Radenovic, J. Brivio, V. Giacometti, A. Kis, Single-layer MoS_2 transistors, Nat. Nanotechnol. 6 (2011) 147-150.

[5] F. Schwierz, Graphene transistors, Nat. Nanotechnol. 5 (2010) 487-496.

[6] L. Li, Y. Yu, G.J. Ye, et al., Black phosphorus field-effect transistors, Nat. Nanotechnol. 9 (2014) 372-377.

[7] S. Haastrup, M. Strange, M. Pandey, et al., The Computational 2D Materials Database, 2D Mater. 5 (2018) 042002.

[8] A. Jain, S.P. Ong, G. Hautier, et al., The Materials Project: A materials genome approach, APL Mater. 1 (2013) 011002.

[9] G.K.H. Madsen, D.J. Singh, BoltzTraP: A code for calculating band-structure dependent quantities, Comput. Phys. Commun. 175 (2006) 67-71.

[10] S. Poncé, E.R. Margine, C. Verdi, F. Giustino, EPW: Electron-phonon coupling and transport, Comput. Phys. Commun. 209 (2016) 116-133.

[11] J. Bardeen, W. Shockley, Deformation potentials and mobilities in non-polar crystals, Phys. Rev. 80 (1950) 72-80.

[12] J. Qiao, X. Kong, Z.-X. Hu, F. Yang, W. Ji, High-mobility transport anisotropy in few-layer black phosphorus, Nat. Commun. 5 (2014) 4475.

[13] J. Schmidt, M.R.G. Marques, S. Botti, M.A.L. Marques, Recent advances in machine learning in solid-state materials science, npj Comput. Mater. 5 (2019) 83.

[14] G.L.W. Hart, T. Mueller, C. Toher, S. Curtarolo, Machine learning for alloys, Nat. Rev. Mater. 6 (2021) 730-755.

[15] Y. Zhuo, A. Mansouri Tehrani, J. Brgoch, Predicting the band gaps of inorganic solids by machine learning, J. Phys. Chem. Lett. 9 (2018) 1668-1673.

[16] L. Ward, A. Dunn, A. Faghaninia, et al., Matminer: An open source toolkit for materials data mining, Comput. Mater. Sci. 152 (2018) 60-69.

[17] C. Chen, Y. Zuo, W. Ye, X. Li, Z. Deng, S.P. Ong, Machine learning of energy materials, Adv. Energy Mater. 10 (2020) 1903242.

[18] H. Cheng, W. Li, L. Tang, et al., From bulk effective mass to 2D carrier mobility via adversarial transfer learning, Nat. Commun. 15 (2024) 5174.

[19] Y. Cai, G. Zhang, Y.-W. Zhang, Polarity-reversed robust carrier mobility in monolayer MoS_2 nanoribbons, J. Am. Chem. Soc. 136 (2014) 6269-6275.

[20] K. Kaasbjerg, K.S. Thygesen, K.W. Jacobsen, Phonon-limited mobility in n-type single-layer MoS_2, Phys. Rev. B 85 (2012) 115317.

[21] eTran2D Database, University of Texas at Austin, https://sites.utexas.edu/yuanyue-liu/etran2d/

[22] M.N. Gjerding, A. Taghizadeh, A. Rasmussen, et al., Recent progress of the C2DB, 2D Mater. 8 (2021) 044002.

[23] Q. Peng, C. Liang, W. Ji, S. De, Mechanical properties of g-GaN, Appl. Phys. A 113 (2013) 483-490.

[24] H. Şahin, S. Cahangirov, M. Topsakal, et al., Monolayer honeycomb structures of group-IV elements, Phys. Rev. B 80 (2009) 155453.

[25] S.M. Lundberg, S.-I. Lee, A unified approach to interpreting model predictions, NeurIPS 30 (2017) 4765-4774.

[26] N.W. Ashcroft, N.D. Mermin, Solid State Physics, Holt, Rinehart and Winston, New York, 1976.

[27] E. Bekaroglu, M. Topsakal, S. Cahangirov, S. Ciraci, First-principles study of defects in silicon carbide honeycomb structures, Phys. Rev. B 81 (2010) 075433.

[28] H.C. Hsueh, G.Y. Guo, S.G. Louie, Excitonic effects in the optical properties of a SiC sheet, Phys. Rev. B 84 (2011) 085404.

[29] Q. Peng, X.-J. Chen, S. Liu, S. De, Mechanical stabilities of graphene-like aluminum nitride, RSC Adv. 3 (2013) 7083-7092.

[30] Z. Zhu, D. Tománek, Semiconducting layered blue phosphorus, Phys. Rev. Lett. 112 (2014) 176802.