Proceedings of the 8th International Conference on Theoretical and Applied Nanoscience and Nanotechnology (TANN 2024)
Chestnut Conference Centre - University of Toronto, Toronto, Canada – June 10-12, 2024
Paper No. 135
DOI: 10.11159/tann24.135

Work Function Tuning for Junctionless Transistor High-K Gate Material Using Machine Learning Descriptor Engineering

Sehtab Hossain[1], Ahmed Fazle Rabbi[2]
[1]University of the Pacific
Atchley Way, Stockton, United State
Shossain1@pacific.edu; Afrabbi@gmail.com
Arlington, Virginia, USA

Abstract **-** A new material descriptor to predict the work function of the high-k gate material of a junctionless transistor. This descriptor model is focused on vectorizing property metrics and empirical property functions along with mixing features. Besides that, with database-based features, the mixing features improve the training and prediction of the models. R2 is greater than 0.89 and the smallest mean absolute error (MAE) and root mean square error (RMSE) are smaller than 0.14 eV and are obtained by implementing Decision Tree Regression. It is noticed that using ensemble algorithms slightly lightens the overfitting problem. This work provides a new method of gate material engineering using material and density functional theory data-based descriptors. Algorithm calculated work function is implemented in Junctionless transistor and performance is observed.

Keywords**:** Junctionless Transistor, FinFET, Descriptor, Inference, Machine Learning

1. Introduction

In the field of nanoelectronics, gate oxide engineering governs the transistor generation. CMOS technology gate oxide used to be silicon dioxide. From 22 nm technology, FinFET is introduced. From FinFET, the gate oxide is changed to Hafnium Dioxide (Hf_2O). With the scaling down process, Hf_2O is infamous for varying work functions. During atomic layer deposition (ALD) because of a very thin layer less than 1 nm, work function variation is noticed with different clusters [1-3]. To tune the work function of Hf_2O, there are several fabrication procedures in existence. But these procedures disregard Density Functional Theory (DFT) calculation. Hence, with the help of Artificial Intelligence (AI), it is possible to pinpoint the crucial reason for work function variation.

Data-driven methods allow resolving the classical methods which demand huge computing resources. Data along with DFT data will cost more - a few hours of computing time. Moreover, it is possible to lessen some computation hours using machine learning algorithms. Using machine learning (ML) to predict the work function of Hf_2O will be a classic illustration of ML in the prediction of physical properties of a material as well as a specific nanoelectronics device. The application of ML and Data Science to material has a lot of avenues to discover new materials and analysis [4,5]. In [4] bi-dimensional (2D) material analysis was done with ML. With ML analysis concentrated on ensemble algorithm band gap and work function are being calculated from a synthetic 2D material database. In [5] 2D MXenes work function was calculated with ML. They introduced reduced order models comprising ten, eight, and five features to predict the work function. They readily tune electronic properties like band gap, work function, density of states at the Fermi level, etc. After running ensemble algorithms, they pointed out the work function. Some authors discuss the integration of machine learning techniques into materials science to expedite the process of discovering new materials. By addressing the challenges inherent in traditional materials discovery methods, such as time-consuming experimentation and limited data analysis capabilities, the authors propose leveraging machine learning to predict material properties, identify potential candidates, and guide experimental efforts [6]. The paper explores various machine learning methodologies, including supervised and unsupervised learning, as well as reinforcement learning, and presents case studies demonstrating successful applications in predicting material properties, optimizing synthesis conditions, and accelerating the discovery of novel materials across different domains. Focusing on the specific target of DUV optical materials, the study employs machine learning models to predict material properties relevant to nonlinear optical applications [7]. Unlike traditional machine learning approaches, the authors emphasize the interpretability of their models, allowing for a deeper understanding of the underlying relationships between

material compositions and desired properties. By leveraging interpretable machine learning, the paper demonstrates an effective strategy for guiding the design process of DUV nonlinear optical materials, ultimately facilitating the discovery of novel materials with tailored properties for advanced optical applications. The authors investigate the effectiveness of machine learning algorithms in predicting materials properties for the purpose of materials discovery [8]. Specifically, the study focuses on assessing the explorative prediction power of these algorithms using k-fold forward cross-validation. The research evaluates a range of machine learning algorithms to determine their ability to accurately predict material properties and uncover potential relationships between input features and desired outcomes. By employing k-fold forward cross-validation, which involves iteratively adding new data points to the training set, the study aims to simulate the exploration process in materials discovery. The findings provide insights into the performance of various machine learning algorithms and their suitability for guiding explorative materials research, offering valuable guidance for researchers seeking to leverage machine learning in the pursuit of new materials discovery.

In this paper, we tune the work function of high-k gate oxide Hf_2O. Hf_2O has several electronic crystal structures with different formation energy. Other material properties like energy above hull eV, magnetic ordering, and DFT properties are considered. Among all crystal structures, the monoclinic electronic crystal structure is widely used for gate oxide. In the database, all these material properties are considered. Along with DFT data, the work function of high-k gate material is predicted. Besides algorithms being implemented to predict work function, descriptor engineering, and feature engineering are crucial for reducing error and loss in the training process to increase the viability of the ML model. Here we make use of Material Project Data [9] of 9900 entries which has reliable data with DFT data. In this paper, descriptors are simple and related to electronic and material properties. To strengthen the efficiency of the model we attached DFT data with electronic and material data. Along with vectorized descriptors, the model could skip expensive-calculated features. Hence, ML cycles will get faster with lesser bias and overfitting. Our findings imply a new insight into material and nanoelectronics domain knowledge. Finally, with ensemble algorithms ML implies some points about important features of material by which work function get affected. From this ML model we could correlate the obtained predicted values with ab initio calculations and determine the work function of high-k gate oxide. After calculating the work function by ML algorithms, the work function related to Hf_2O is implemented in junctionless transistor, and performance is analyzed. This paper is organized as follows: section 1 discusses the introduction where related works are described. Section 2 is dedicated to methodology where several ML algorithms and their implementations are discussed elaborately. Section 3 is dedicated to results and discussion and ends with a conclusion.

2. Methodology

2.1. Descriptor Construction

To construct a descriptor, we introduce relevant atom-atom interactions based on elemental parameters to produce molecule-level descriptors. In this study, in total eleven descriptors are implemented in this paper. Among them, crystal systems, space group symbols, volume, density, and sites are electronic-crystallographic data. One descriptor Energy Above Hull (EAH) in ev/atom relates to assessing the stability of a material concerning its possible decomposition products. Specifically, it quantifies the energy difference between the material in question and the most stable composition of its constituent elements. When a material is synthesized or computed, it may not always be the most stable configuration given its constituent elements. The energy above hull is calculated by comparing the total energy of the material to the total energy of the most stable combination of its constituent elements, typically under the same conditions (such as pressure and temperature). A positive value of energy above hull indicates that the material is energetically less stable compared to its constituent elements in their most stable forms. Another descriptor formation energy in eV/atom represents the energy change associated with the formation of a compound from its constituent elements, per atom. It is defined as the energy difference between the total energy of the compound and the total energies of its constituent elements in their stable reference states. Mathematically, it is often expressed as

$$Ef = E_{compound} - \sum_i n_i \, Eelement_i \qquad (1)$$

where Ef is the formation energy, $E_{compound}$ the total energy of the compound, n_i is the number of atoms of an element in the compound and $Eelement_i$ is the total energy of one atom of the element is in its stable reference state. A negative formation energy indicates that the compound is energetically stable concerning its constituent elements. This means that it requires less energy to form the compound from its elements than to keep them separated. Conversely, positive formation energy suggests that the compound is energetically unfavorable and may tend to decompose into its constituent elements. After selecting descriptors, data preparation, data augmentation, and feature engineering commenced. The data frame is cross-validated and applied to the ML algorithm.

2.2. Model Selection

The database has been implemented to several ML algorithms ranging from Linear, and Elasticnet to ensemble algorithms like Random Forest and Extreme Gradient Boost (XGB). After data cleaning and preparation and feature engineering, the dataset was split into 1:4 test to train ratio respectively. For each ML model, datasets are shuffled and analyzed with tuned Hyperparameters. Hyperparameters are tuned with GridSearchCV of scikit-learn. The best model is selected from the lowest function loss and R_2 score. Cross-validation is applied in the train-test set to reduce the overfitting of data. The cross-validation sampling is generated by k-fold cross-validation for regression analysis. Moreover, after random splits the selected models are stable.

3. Result and Discussion

The outcome of the evaluation is based on features. The dataset has 3 hybrid features and 8 reference features. the data set is at first prepared and feature-engineered and then split into train and test sets (Fig.1). Before splitting exploratory data analysis is done meticulously.

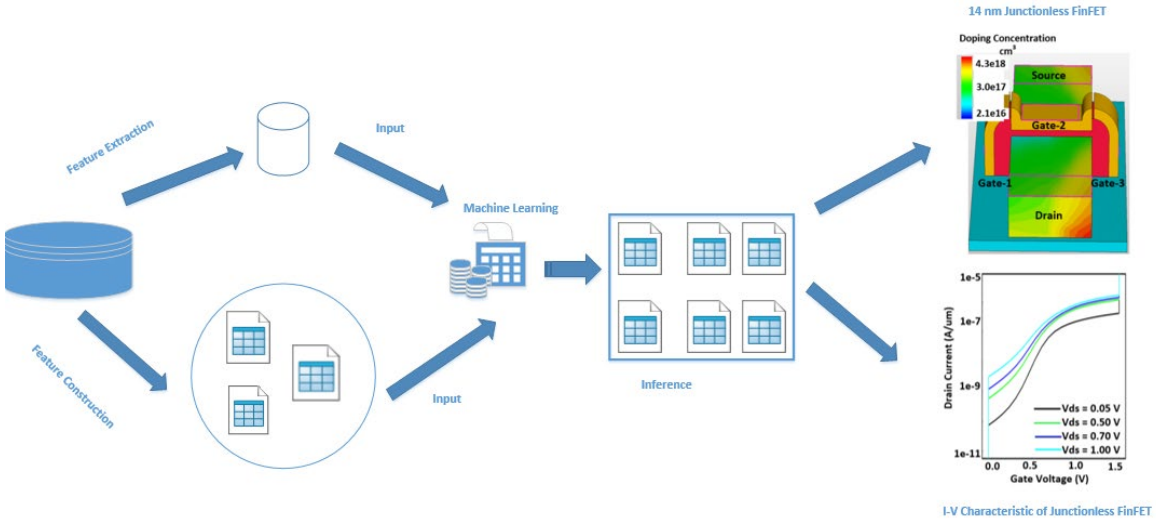

Fig 1. Workflow of the descriptor engineering and ML algorithm implementation. After extracting features and constructing vectorized data from augmentation implemented as input to machine learning algorithms and inferenced data is utilized to build 14 nm technology Junctionless FinFET.

3.1 Exploratory Data Analysis (EDA)

For exploratory data analysis, correlation analysis is implied through features. In correlation analysis, all the features exhibit Pearson correlation is less than 0.49. Hence, in correlation analysis, the features are not strongly correlated. To investigate the strong relation between features, the Variation Inflation Factor (VIF) is calculated. After VIF calculation, it is noticed that features 'sites' and 'volume' have high VIF scores which implies that these two features are strongly correlated (dependent on each other). From pair plot analysis, bandgap and work function show cluster behaviour. Energy Above hull (EAH) and bandgap also exhibit cluster behaviour.

3.2 Regression Analysis

Several regression algorithms are implemented in this dataset. Linear regression, Ridge regression, Lasso Regression, and Elastic Net regression show almost the same inference. Linear regression algorithms have higher Mean Absolute Error (MAE) and Root Mean Square error (RMSE) than other algorithms. R^2 Score is almost the same. MAE, RMSE, and R^2 Score infer the nonlinear behaviour of the dataset.

3.3 Support Vector Machine & K Nearest Neighbor (KNN) Algorithm

In Support Vector Regression (SVR) MAE, RMSE reduced significantly, and R2 Score (0.82) improved sharply. Radial Base Function (RBF) kernel is implemented with C value of 10 and gamma of 'scale'. KNN regression shows slightly higher MAE and RMSE than SVR but R^2 Score is slightly higher (0.84). The variation of R^2 Score may be the performance difference between RBF kernel for SVR and the number of neighbors (7) for KNN.

3.4 Neural Network Regression

With 'Adam' optimizer and 'mean Squared Logarithmic Error' metrics neural network regression provides MAE, RMSE of 0.14 eV and 0.14 eV respectively. Neural network optimized with Adam optimizer and 'relu' activation, is the reason for low MAE and RMSE value. R^2 Score for the train set and test set are 0.87 and 0.84 respectively. A slight difference between these two sets may be caused by overfitting. Fig.2 depicts Model loss and Model Mean Squared logarithmic error. Observing Fig.2, it is clear that with the increment of epoch train and test set error is decreasing, and the error difference is decreasing.

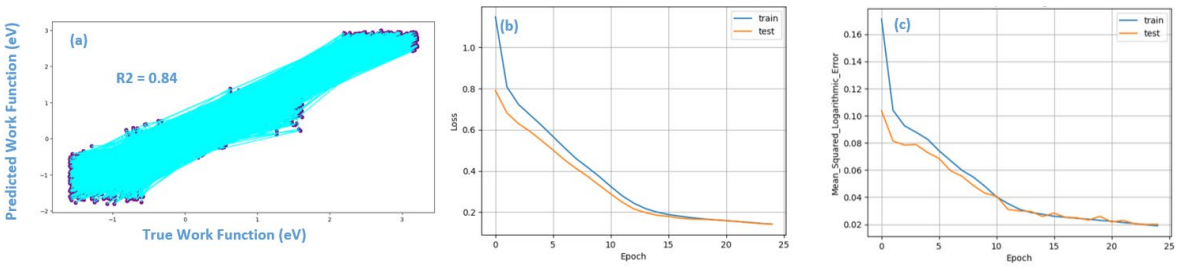

Fig 2. (a) Visualization of the predicted and true values of the work function with Neural network Regression (b) Error versus Epoch of Neural network (c) Mean squared logarithmic error vs epoch, for both loss function decrease with the epoch of time.

3.5 Tree Algorithm

Tree algorithm- decision tree regression shows some crucial points. The feature importances plot shows the feature 'band gap' as the highest importance than other features (Fig.3(c)). RMSE and MAE are in the limit of 0.3 eV and a slight difference in R^2 Score is observed between train and test set. In Fig. 3 (a,b), for the case of the test set there are some small outliers visible.

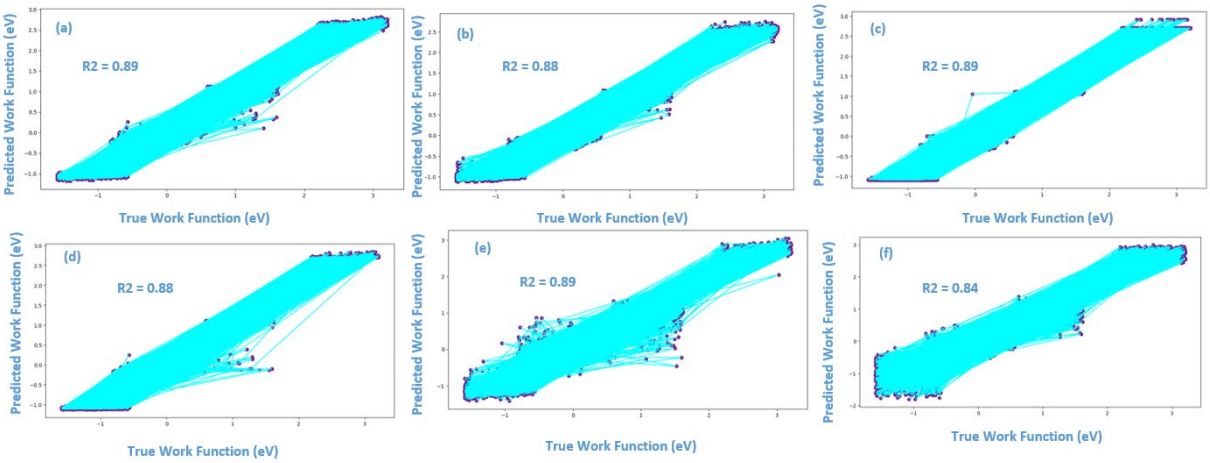

Fig 3. Visualization of predicted and true values of work function with several ML algorithms (a) Gradient Boost Regression (b) XGBoost Regression (c) Decision Tree Regression (d) Random Forest Regression (e) Neural Network Regression (f) Support Vector Regression. R^2 Score is also mentioned in the images.

3.6 Ensemble Algorithm

Random forest and Extreme Gradient Boost (XGB) are implied as ensemble algorithms. Apart from decision tree regression, these two ensemble algorithms imply 'density' as the highest importance feature. MAE and RMSE are in the range of 0.3 eV and both of the algorithms show differences in MAE, RMSE, and R^2 Score (Fig 3(b, c, d)). Both algorithms show R^2 score in the range of 0.88 and a little bit of difference in the train and test set (Fig.4). Analyzing feature importance, it has been noticed that tree algorithms and all ensemble algorithms show 'density' and next 'band gap' as the highest important features in the models (Fig. 5).

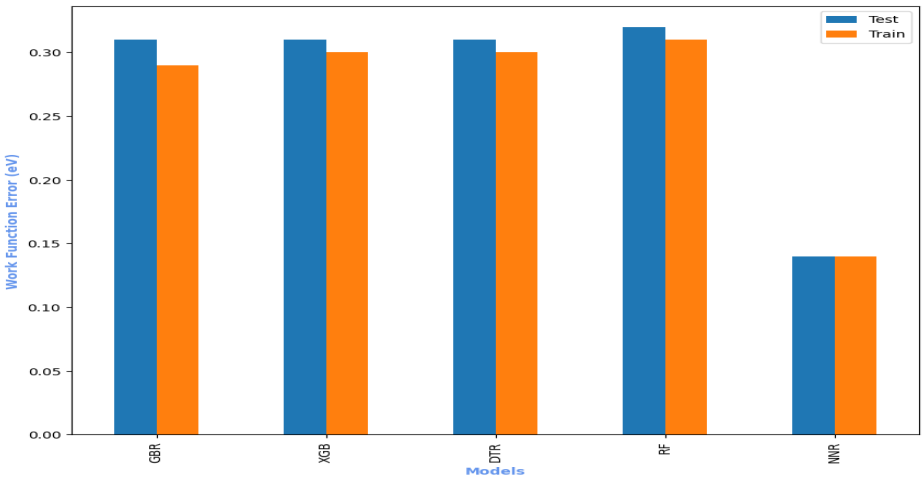

Fig 4. Bar plot of the RMSE errors generated by several ML Algorithms. The bars are grouped by train and test sets.

After several inferences, the calculated work function is implemented in a 14 nm technology junctionless FinFET transistor by Sentaurus TCAD, and i-v characteristics are observed. From the inferenced work function, the transistor's electron transport is normal and On current is at the scale of 1e -6 A [11-13] (Fig.6(a, b)). Additionally, logic states are implemented in the FinFET, and electron transport and ON current are observed and on current is in very good agreement with concurrent FinFET devices and pointed out a little bit max current than concurrent FinFET devices (Fig. 6 (c,d)). is examined in several switching conditions for further dimensions of the transistor and noticed On current is in the scale 1exp-6 A.

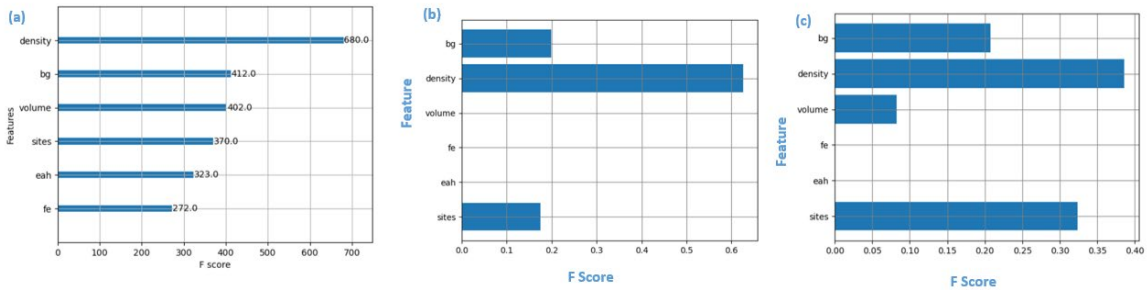

Fig 5. Feature importance in several ensemble algorithms (a) XGB (b) DTR (c) RF Where bg = bandgap, eah = Energy Above Hull, fe = Formation energy

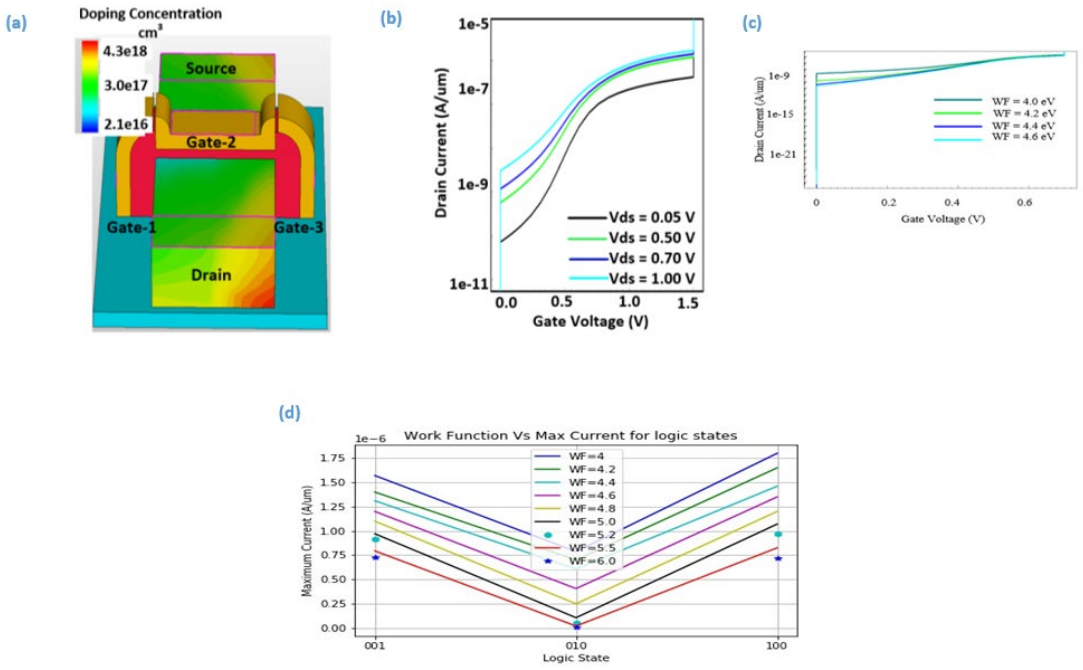

Fig 6. Implementation of the inferenced work function in a tri-gate junctionless transistor. (a) Trigate junctionless 14 nm technology FinFET transistor where inferenced value of work function = 4.66 eV is implemented (b) I-V characteristic – performance of 14 nm FinFET (c) Drain current variation with varying inferenced work function (d) work function Vs maximum current varying several logic states in the FinFET.

4. Conclusion

In summary, vectorized descriptors along with material data descriptors are implemented for the proposed model to infer the work function of high-k gate oxide material. We show that the whole pipeline of the model is low-cost and highly efficient. After the implementation of the database in the model ensemble algorithms show a high R^2 value (>0.89) and MAE and RMSE at a maximum of 0.3 eV. Neural network and SVR show great improvement with the lowest MAE and RMSE (> 0.14 eV) and decent R^2 value (>0.89). From this observation, ii is evidence of the database nonlinearity. Inference from ensemble algorithm and feature importance is in good agreement with empirical data [15,16]. Furthermore, inferences from ML models are meticulously observed to control overfitting and bias. Finally, the inference work function is applied to a 14nm technology junctionless FinFET transistor and the performance is the same mark as the existing junctionless transistor. This work provides a pipeline for ML-led nanoelectronics with low resource requirements and high efficiency.

References

[1] Sampson, Jenyfal, Sivakumar P, and Velmurugan SP. "Effect of work Function Tuning over the Device Characteristics of GAA SNSTFT." Silicon 15.7 (2023): 3201-3214.

[2] Fields, S. S., Jaszewski, S. T., Lenox, M. K., & Ihlefeld, J. F. (2023). Asymmetric Electrode Work Function Customization via Top Electrode Replacement in Ferroelectric and Field-Induced Ferroelectric Hafnium Zirconium Oxide Thin Films. Advanced Materials Interfaces, 10(8), 2202232.

[3] Schaeffer, J. K., Fonseca, L. R. C., Samavedam, S. B., Liang, Y., Tobin, P. J., & White, B. E. (2004). Contributions to the effective work function of platinum on hafnium dioxide. Applied physics letters, 85(10), 1826-1828.

[4] Smith, B. "An approach to graphs of linear forms (Unpublished work style)." (1995).

[5] Roy, P., Rekhi, L., Koh, S. W., Li, H., & Choksi, T. S. (2023). Predicting the work function of 2D MXenes using machine-learning methods. Journal of Physics: Energy, 5(3), 034005.

[6] Juan, Y., Dai, Y., Yang, Y., & Zhang, J. (2021). Accelerating materials discovery using machine learning. Journal of Materials Science & Technology, 79, 178-190.

[7] Wu, M., Tikhonov, E., Tudi, A., Kruglov, I., Hou, X., Xie, C., ... & Yang, Z. (2023). Target-Driven Design of Deep-UV Nonlinear Optical Materials via Interpretable Machine Learning. Advanced Materials, 35(23), 2300848.

[8] Xiong, Z., Cui, Y., Liu, Z., Zhao, Y., Hu, M., & Hu, J. (2020). Evaluating explorative prediction power of machine learning algorithms for materials discovery using k-fold forward cross-validation. Computational Materials Science, 171, 109203.

[9] Anubhav Jain, Shyue Ping Ong, Geoffroy Hautier, Wei Chen, William Davidson Richards, Stephen Dacek, Shreyas Cholia, Dan Gunter, David Skinner, Gerbrand Ceder, Kristin A. Persson; Commentary: The Materials Project: A materials genome approach to accelerating materials innovation. APL Mater. 1 July 2013; 1 (1): 011002. https://doi.org/10.1063/1.4812323

[10] https://www.synopsys.com/manufacturing/tcad.html

[11] Hossain, Sehtab. Function Implementation in a Multi-Gate Junctionless Fet Structure. Diss. University of Missouri-Kansas City, 2023.

[12] Hossain, Sehtab, Md Arif Iqbal, and Mostafizur Rahman. "A new approach towards embedded logic in a single device." 2020 IEEE 20th International Conference on Nanotechnology (IEEE-NANO). IEEE, 2020.

[13] Hossain, S., Iqbal, M. A., Samant, P., Siddiki, M. K., & Rahman, M. (2023). More Than a Device: Function Implementation in a Multi-Gate Junctionless FET Structure. Journal of Electronics and Electrical Engineering, 1-11.

[14] Hossain, Sehtab. "Iridium Modified Silicon (001) Surface." (2015).

[15] Fatima, F., Can Oguz, I., Çakır, D., Hossain, S., Mohottige, R., Gulseren, O., & Oncel, N. (2016). On the structural and electronic properties of Ir-silicide nanowires on Si (001) surface. Journal of Applied Physics, 120(9).

[16] Wei, Jing, Xuan Chu, Xiang-Yu Sun, Kun Xu, Hui-Xiong Deng, Jigen Chen, Zhongming Wei, and Ming Lei. "Machine learning in materials science." InfoMat 1, no. 3 (2019): 338-358.

[17] Morgan, Dane, and Ryan Jacobs. "Opportunities and challenges for machine learning in materials science." Annual Review of Materials Research 50 (2020): 71-103.
[18] Zhang, Ying, and Chen Ling. "A strategy to apply machine learning to small datasets in materials science." Npj Computational Materials 4.1 (2018): 25.
[19] Schleder, G. R., Padilha, A. C., Acosta, C. M., Costa, M., & Fazzio, A. (2019). From DFT to machine learning: recent approaches to materials science–a review. Journal of Physics: Materials, 2(3), 032001.
[20] Liu, G., Sohn, S., O'Hern, C. S., Gilbert, A. C., & Schroers, J. (2024). Effective subgrouping enhances machine learning prediction in complex materials science phenomena: Inoue's subgrouping in discovering bulk metallic glasses. Acta Materialia, 265, 119590.
[21] Ye, Z., Li, J., Wang, W., Qin, F., Li, K., Tan, H., & Zhang, C. (2024). Data-driven visualization of the dynamics of machine learning in materials research. Journal of Cleaner Production, 141410.
[22] Lian, Z., Ma, Y., Li, M., Lu, W., & Zhou, W. (2024). Discovery Precision: An effective metric for evaluating performance of machine learning model for explorative materials discovery. Computational Materials Science, 233, 112738.
[23] Long, C., Liu, S., Sun, R., & Lu, J. (2024). Impact of structural characteristics on thermal conductivity of foam structures revealed with machine learning. Computational Materials Science, 237, 112898.
[24] Nelaturu, P., Hattrick-Simpers, J. R., Moorehead, M., Jambur, V., Szlufarska, I., Couet, A., & Thoma, D. J. (2024). Multi-principal element alloy discovery using directed energy deposition and machine learning. Materials Science and Engineering: A, 891, 145945.
[25] Chen, Ming, Zhenhua Yin, Zhicheng Shan, Xiaokai Zheng, Lei Liu, Zhonghua Dai, Jun Zhang, Shengzhong Frank Liu, and Zhuo Xu. "Application of machine learning in perovskite materials and devices: a review." Journal of Energy Chemistry (2024)
[26] Li, He, Zechen Tang, Jingheng Fu, Wen-Han Dong, Nianlong Zou, Xiaoxun Gong, Wenhui Duan, and Yong Xu. "Deep-learning density functional perturbation theory." Physical Review Letters 132, no. 9 (2024): 096401.