**Integrating the Asymptotic Iteration Method with Machine Learning for Predicting Vibrational Energy Levels of Diatomic Molecules**

**ABSTRACT**

|  |
| --- |
| **Aim:** This study integrates the Asymptotic Iteration Method (AIM) with Machine Learning (ML) models to enhance the prediction of vibrational energy levels in diatomic molecules. Traditional quantum mechanical methods, while accurate, are computationally demanding. This study aims to determine whether ML models can approximate these calculations efficiently while maintaining high accuracy.  **Methodology: The v**ibrational energy levels of Li₂, CN, and CO molecules were computed using AIM within the Morse potential framework. Three ML models—Random Forest (RF), Gradient Boosting (GB), and Support Vector Regression (SVR)—were trained using AIM-derived datasets. The models were evaluated using **Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and R² score** to assess their predictive performance.  **Results:** SVR demonstrated the highest predictive accuracy, achieving a R² score of **0.999650**, the lowest MAE and RMSE values of **0.124391** and **0.158412 respectively**, outperforming RF and GB. The results indicate that ML models, particularly SVR, can effectively approximate AIM calculations with minimal error. Furthermore, **3D potential energy surface visualizations confirmed the strong agreement between ML and AIM predictions**, validating the reliability of ML-based approaches.  **Conclusions:** This study demonstrates that **ML can serve as an efficient and scalable alternative** to traditional quantum mechanical methods for predicting vibrational energy levels. The findings have implications for computational chemistry, spectroscopy, and materials science by **reducing reliance on computationally intensive calculations**. However, the study is limited by **data generalization**, as accuracy depends on the diversity of the training dataset. Future work should focus on **expanding datasets, integrating deep learning techniques, and exploring hybrid AIM-ML models** to improve generalizability and predictive robustness. |

**Keywords:** Machine Learning, Asymptotic Iteration Method, Vibrational Energy, Diatomic Molecules

**1.0 INTRODUCTION**

Diatomic molecules, consisting of two atoms, exhibit quantized vibrational energy levels governed by the Schrödinger equation. The Morse potential, introduced as a model for describing vibrational energy in such systems (Omoriwhovo et al., 2022), has become fundamental across multiple disciplines, including molecular physics, chemical physics, quantum chemistry, materials science, and solid-state physics (Barakat & Abodayeh, 2006; Yang et al., 2022). Traditionally, vibrational energy levels are calculated using first-principles or semi-empirical quantum mechanical approaches, such as Hartree-Fock theory, density functional theory (DFT), perturbation theory, vibrational self-consistent field (VSCF), vibrational configuration interaction (VCI), vibrational coupled cluster (VCC) theory, variational methods, supersymmetric quantum mechanics, and factorization methods (Christiansen, 2007; Omoriwhovo et al., 2022). While these methods provide an elaborate theoretical framework, their high computational cost and poor scalability with increasing system size pose significant challenges, particularly for complex systems (Jensen, 2017).

Recent advancements in machine learning (ML) offer a promising alternative to accelerate quantum chemical calculations and predict molecular properties with high accuracy. Studies indicate that ML models can achieve predictive accuracies comparable to or even surpassing those of DFT (Faber et al., 2017). Various ML methodologies, including kernel ridge regression, random forests, and neural networks, have been employed to predict electronic and energetic properties of organic molecules, utilizing molecular representations such as Coulomb matrices, bag-of-bonds, and graph-based methods (Hansen et al., 2015; Faber et al., 2017). These approaches use large datasets and sophisticated algorithms to approximate solutions to the Schrödinger equation without the need for direct computation, significantly reducing computational costs (Manzhos, 2020). However, challenges remain regarding the transferability of molecular representations, model generalization across materials, and ensuring reproducibility (Rupp, 2015; Aldossary et al., 2024).

Despite these advancements, ML models often lack physical interpretation and generalization due to their reliance on empirical features. Many existing models function as black-box systems that do not explicitly incorporate fundamental quantum mechanical principles (Bikmukhametov & Jäschke, 2020). This limitation underscores the necessity of integrating physics-based features into ML frameworks to enhance both interpretability and predictive accuracy.

One promising approach involves incorporating the Asymptotic Iteration Method (AIM) as a source of physically meaningful molecular descriptors to improve ML-based predictions of vibrational energy levels. Within the Morse potential framework, AIM has demonstrated effectiveness in characterizing vibrational energy spectra for diatomic molecules (Apanavičius et al., 2021; Omoriwhovo et al., 2022). In a hybrid AIM-ML framework, ab initio calculations provide high-quality training data and fundamental physical insights, while ML models refine these predictions by identifying subtle patterns and correlations in the data (Huang & von Lilienfeld, 2021). This integration mitigates limitations such as poor extrapolation and restrictive spatial locality assumptions that commonly affects conventional ML approaches (Fedik et al., 2023). The development of physics-informed structural representations is critical for constructing efficient and universally applicable models for predicting molecular properties (Musil et al., 2021).

Hybrid modeling, physics-informed ML, and model calibration techniques further enhance accuracy and efficiency by embedding mechanistic information within data-driven models (Bradley et al., 2022). These methodologies have been successfully employed in diverse areas of chemistry, including the construction of potential energy surfaces, molecular dynamics simulations, and property prediction. The fusion of ML and ab initio techniques has significantly advanced computational catalysis, overcoming the limitations of traditional molecular simulations for complex systems (Orupattur et al., 2020). This study extends these hybrid approaches to the prediction of vibrational energy levels, investigating whether AIM-derived parameters can refine ML-based energy level predictions for diatomic molecules.

**2.1 THEORETICAL FRAMEWORK**

**2.1.1 Numerical formalism of the Asymptotic Iteration Method**

The Morse potential have been one of the most helpful and convenient models and this provides a qualitative description of the interaction between two atoms in a molecule. From analysis, it has been shown that the rotational energy of a molecule is far smaller than the vibrational energy and so, in a very pure Morse potential model the rotational energy of a molecule is being omitted (Barakat and Abodayeh,2006; Gyuk et al, 2019; Omoriwhovo et al., 2022). Thus, we begin with the Schrodinger equation for one-dimension given by;

(2.1)

where E represents the energy eigenvalues and  is the Morse potential function in one dimension given as;

(2.2)

where is the dissociation energy, xe is the equilibrium inter-nuclear distance of a diatomic molecule, is the reduced mass and is an adjustable parameter. The Morse potential has a minimum value at and it asymptotically approaches De as .

To solve the Schrodinger equation, we introduce the Morse’s substitution, i.e. ,

The Schrodinger equation in equation (2.1) becomes

(2.3)

Rewriting the derivatives of equation (2.3) using , we obtain

(2.4)

Expanding the term and dividing through by , we obtain

(2.5)

Let

Equation (2.5) can be reexpressed as

(2.6)

Equation (2.6) can be solved by introducing the ansatz

(2.7)

This results to,

(2.8)

where s is a parameter to ensure the correct asymptotic behavior and is related to the energy.

To solve equation (2.3) we express it in the general AIM form

(2.9)

where

(2.10)

On differentiating both sides of equation (2.5) with respect to repeatedly, a recursion expression is obtained . This is given as

(2.11)

For sufficiently large n, the asymptotic method requires that

(2.12)

Equation (2.12) provides the quantization condition for solving the eigenvalues iteratively.

**2.1.2 Analytical Formalism of the Vibrational Energy levels using AIM**

The energy eigenvalues (En) for the Schrodinger equation with the Morse potential were calculated using the equation (2.12). To generate this energy spectrum En, the equation (2.12) is first solved for En where the iteration should be terminated by applying a condition as a rough calculation to equation (2.12).

On the other hand, for each iteration, the expression

, (2.13)

where the expression above depends on En and . The calculated En by means of this condition should, however be independent of the choice of . Nevertheless, the choice of is observed to be critical only to speed the convergence to En, as well as for the stability of the process.

In this study, our best starting value for is the value at which the effective potential of equation (2.9) takes its minimum value, that is when =1 (Barakat and Abodayeh, 2006; Gyuk et al, 2019; Omoriwhovo et al., 2022). Thus, at the end of the iterations, =1, the generalized results of the AIM for En for is given as

(2.14)

For different values of n the equation (2.14) can be given for n=0,1,2,3…;

(2.14a)

(2.14b)

(2.14c)

Thus, the exact energy eigenvalues of the Morse potential are;

(2.15)

where En is the vibrational energy levels of the system.

**3.0 METHODOLOGY**

This study employed a systematic approach to assess the effectiveness of machine learning (ML) models in predicting the vibrational energy levels of diatomic molecules. The methodology is structured into distinct stages, including data collection and preprocessing, model selection, evaluation, and implementation.

**3.1 Data Collection and Preprocessing**

The vibrational energy levels of three diatomic molecules were computed using the Asymptotic Iteration Method (AIM) within the Morse potential framework. The Schrödinger equation was solved for each molecule to derive the potential energy surface (PES), from which vibrational energy levels were extracted using the harmonic oscillator approximation. The dataset comprises diatomic molecules in various electronic states, as detailed in Table 1 (Majedifar & Islampour, 2012). Prior to model training, the dataset was cleaned by removing outliers and inconsistencies, and missing values were imputed using regression techniques. The data was subsequently partitioned into training and test sets, with an 80:20 split to ensure robust model evaluation.

Table 1 Basic molecular data used for the numerical computation of the energy eigenvalues of the Morse potential (Majedifar and Islampour, 2012)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Molecule | Electronic State |  |  |  |  |
|  |  | 8541 | 2.67328 | 0.867 | 351.43 |
|  | 8434.5 | 2.936 | 0.6698 | 270.12 |
|  | 8940 | 3.10821 | 0.616 | 255.47 |
| CN |  | 76600 | 1.17182 | 2.32 | 2068.59 |
|  | 55200 | 1.2333 | 2.30 | 1812.5 |
|  | 53900 | 1.15 | 2.88 | 2163.9 |
| CO |  | 88575.905 | 1.128323 | 2.3423473 | 2169.813 |
|  | 29704.287 | 1.2353 | 2.830202 | 1518.2 |

**3.2 Machine Learning Models**

A selection of ML models was employed to compare predictive performance across different algorithmic approaches. Linear Regression (LR) was used as a baseline model to establish a linear relationship between molecular parameters and vibrational energy levels (Fransson et al., 2019). Support Vector Machines (SVM), incorporating non-linear kernels such as the radial basis function, were applied to capture complex data relationships (Du et al., 2024; Mohammed & Mahdi, 2024). Gradient Boosting Regressors, an ensemble learning method, were also implemented to aggregate multiple weak learners (typically decision trees) into a strong predictive framework (Fafalios et al., 2020; Chintala, 2019). Each model was trained using the preprocessed dataset, with hyperparameter tuning performed via a validation set.

**3.3 Model Evaluation**

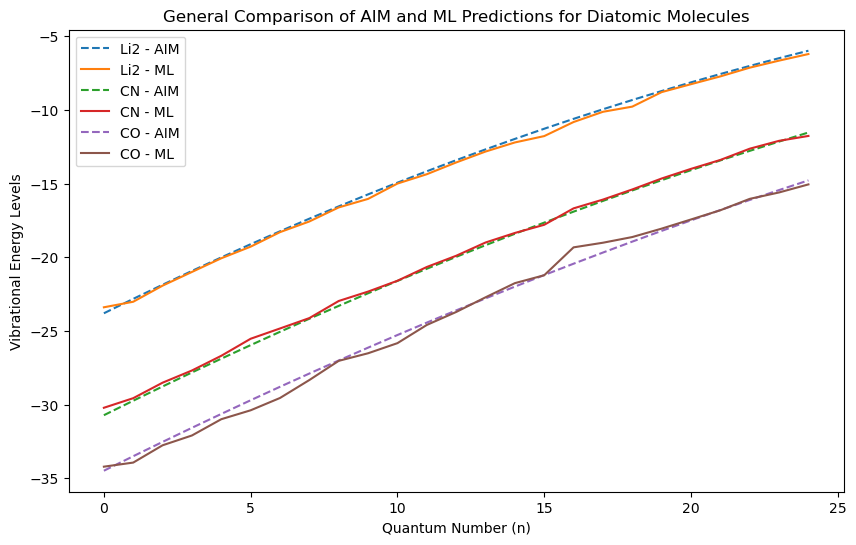
The predictive performance of each machine learning (ML) model was evaluated using **Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and the R² score**. These metrics collectively assess the **accuracy and reliability** of the models, with MAE quantifying the average absolute deviation between predicted and actual vibrational energy levels, RMSE captures the magnitude of prediction errors, and the R² score measures the proportion of variance explained by the model.

**3.4 Implementation**

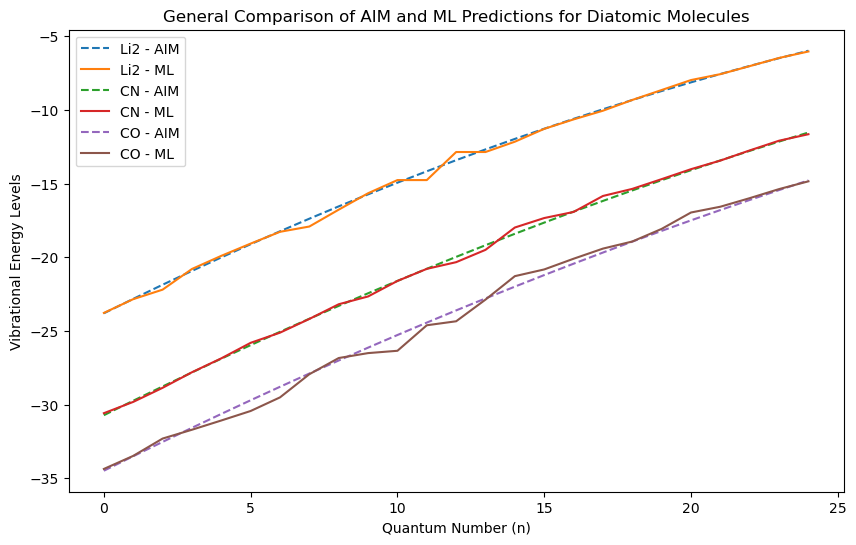
The ML models were implemented using Python and its associated scientific computing libraries, such as scikit-learn. While accuracy, reliability, and computational speed are crucial in model selection, accessibility and ease of use of the programming language remain key considerations in emerging economies (Akpojotor et al., 2010; Akpojotor and Ehwerhemuepha, 2012).

**4.0 RESULTS AND DISCUSSION**

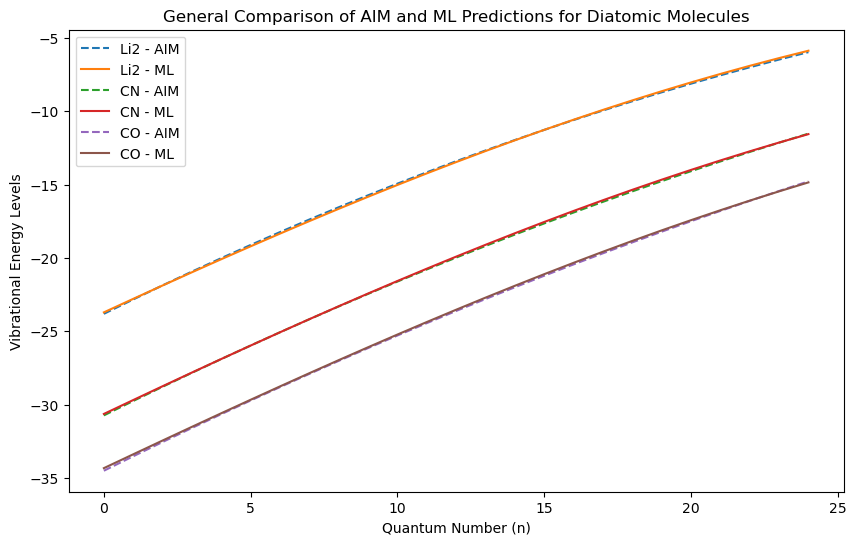
The study presents an integration of the Asymptotic Iteration Method (AIM) with Machine Learning (ML) models—Random Forest (RF), Gradient Boosting (GB), and Support Vector Regression (SVR)—for predicting the vibrational energy levels of diatomic molecules (Li₂, CN, and CO). Figures 1–3 illustrate the comparison between AIM-computed and ML-predicted vibrational energy levels for Li₂, CN, and CO. The results obtained from the ML models agrees closely with the AIM derived results with high precision which supports the hypothesis that data-driven methods can approximate physics-based numerical approaches when trained on sufficiently rich datasets  (Sturm et al., 2020; Jamali et al., 2021). The ML approximation closely aligns with the AIM results obtained by Omoriwhovo et al. (2022) using the framework of Morse potential to obtain the vibrational energy spectra of diatomic molecules.



**Figure 1: Comparison of AIM and ML Predictions for Diatomic Molecule using Random Forest Regression**



**Figure 2: Comparison of AIM and ML Predictions for Diatomic Molecule using Gradient Boosting**



**Figure 3: Comparison of AIM and ML Predictions for Diatomic Molecule using Support Vector Machine**

The ability of ML models to align with the AIM predictions lies on the quality and diversity of the input data. The input features—Dissociation Energy (D), Equilibrium Bond Length (x₀), Anharmonicity Constant (γ), and Vibrational Frequency (w)—are the fundamental molecular parameters that define vibrational behavior. The effectiveness of MinMax scaling in normalizing these features ensures that the models generalize well across different molecular systems.

Recent studies have emphasized the importance of feature engineering in machine learning applications for quantum mechanics ( Janet & Kulik, 2017; Dral, 2020; Zhang et al., 2023).The present work aligns with these findings, demonstrating that proper feature normalization and selection significantly improve ML model performance in molecular property prediction.

Table 2 presents the Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and R² Score for each ML model emphasising on the model performance comparison for predicting the vibrational energy of the diatomic molecules.The results obtained indicated that SVR outperforms the other models,with the lowest MAE and RMSE values of 0.124391 and 0.158412 respectively and the highest R² score of 0.999650.This suggests a near-perfect alignment with AIM predictions as observed in studies by Omoriwhovo et al.(2022).

**Table 2: Model Preformace Comparison for Predicting the Vibrational Energy of Diatomic Molecules**

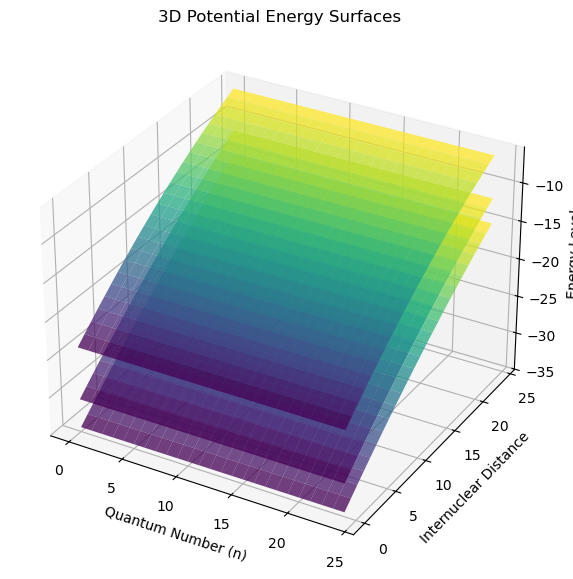
|  |  |  |  |
| --- | --- | --- | --- |
| Model | Mean Absolute Error (MAE) | Root Mean Square Error (RMSE) | R2 Score |
| Random Forest | 0.653578 | 0.790309 | 0.991288 |
| Gradient Boosting | 0.383333 | 0.485214 | 0.996716 |
| Support Vector Regression | 0.124391 | 0.158412 | 0.999650 |

The out performance of SVR model can be attributed to its ability to effectively transform nonlinear relationships into linear through kernel-based methods (Dhhan et al., 2018). Recent studies have demonstrated the effectiveness of support vector regression (SVR) in modeling various chemical and physical properties with mininal error (Nguyen, 2019;Tawfik et al., 2019)., making it an ideal choice for vibrational energy level predictions.

Gradient Boosting (GB), while slightly less accurate than SVR, still provides a strong predictive performance. This is consistent with its well-known advantages capturing complex dependencies and interactions among variables, even in high-dimensional datasets (Zhang et al., 2019).  Random Forest (RF), while still accurate, exhibits higher MAE and RMSE, likely due to its ensemble nature, which may not fully capture the smooth functional relationship between molecular parameters and energy levels.

Figure 4 depicts the 3D visualizations of potential energy surfaces (PES) for the diatomic molecules. These plots show how energy levels change with internuclear distance, providing understanding about the molecular stability and anharmonicity effects of the diatomic system.

The Random Forest, Gradient Boosting and SVR models were successfully in reproducing the curvature of AIM-derived energy surfaces, indicating their potential as reliable approximators of physical energy landscapes. Similar ML-driven potential energy surface reconstructions have been reported in the literature  (Han et al., 2021; Tao et al., 2022)., reinforcing the role of ML in accelerating quantum mechanical calculations and predictions for molecular system



**Figure 4: 3D Potential Energy Surface of the Diatomic Molecules**

**5.0 CONCLUSION**

This study have successfully integrated the Asymptotic Iteration Method (AIM) with Machine Learning (ML) models to predict the vibrational energy levels of diatomic molecules. The results obtained demonstrated that ML models, particularly the **SVR model**, achieved the highest accuracy, with a **R² score of 0.999650 which closely aligned** with AIM predictions. The results confirms that ML can serve as an efficient and low cost computational alternative to traditional quantum mechanical methods. The broader implications of this study extend to **accelerating quantum chemistry and quantum mechanical calculations**, reducing reliance on high-cost **ab initio simulations**, and enabling **real-time predictions of molecular properties** in larger chemical and material systems. The findings have **practical applications** in spectroscopy, materials science, and computational chemistry, where rapid and reliable vibrational energy estimations are important. However, the study is limited by **model generalizability**, as predictions remain dependent on training data quality and molecular diversity. Future research should focus on **using larger datasets**, incorporating **deep learning approaches**, and **exploring hybrid AIM-ML frameworks** to enhance model robustness and extend applications to more complex molecular systems.

**Disclaimer (Artificial intelligence):**

**Author(s) hereby declare that NO generative AI technologies such as Large Language Models (ChatGPT, COPILOT, etc.) and text-to-image generators have been used during the writing or editing of this manuscript. All script used for the model comparison will be provided on request**

**REFERENCES**

Omoriwhovo, J.O., Agbajor, G.K.,Akpolilie, A.L. & Akpojotor, G. (2022). Python AIM solver to Obtain Vibrational Energy Spectra of Diatomic Molecules. African Journal of Physics, 15:15-31

Barakat, T. and Abodayeh, K. (2006). Exact Solutions for Vibrational Levels of the Morse Potential via the Asymptotic iteration Method. Czechoslovak Journal of Physics. 56:583-590

Yang, Z., Wan, Z., Liu, L., Fu, J., Fan, Q., Xie, F. et al. (2022). Achieving vibrational energies of diatomic systems with high quality by machine learning improved DFT method. RSC advances, 12(55), 35950-35958.

Christiansen, O. (2007). Vibrational structure theory: new vibrational wave function methods for calculation of anharmonic vibrational energies and vibrational contributions to molecular properties. Physical Chemistry Chemical Physics, 9(23), 2942-2953.

Jensen, F. (2017). Introduction to computational chemistry. John wiley & sons.

Faber, F. A., Hutchison, L., Huang, B., Gilmer, J., Schoenholz, S. S., Dahl, G. E., ... & Von Lilienfeld, O. A. (2017). Prediction errors of molecular machine learning models lower than hybrid DFT error. Journal of chemical theory and computation, 13(11), 5255-5264.

Hansen, K., Biegler, F., Ramakrishnan, R., Pronobis, W., Von Lilienfeld, O. A., Muller, K. R., & Tkatchenko, A. (2015). Machine learning predictions of molecular properties: Accurate many-body potentials and nonlocality in chemical space. The journal of physical chemistry letters, 6(12), 2326-2331.

Manzhos, S. (2020). Machine learning for the solution of the Schrödinger equation. Machine Learning: Science and Technology, 1(1), 013002.

Rupp, M. (2015). Special issue on machine learning and quantum mechanics. International Journal of Quantum Chemistry, 115(16), 1003-1004.

Aldossary, A., Campos‐Gonzalez‐Angulo, J. A., Pablo‐García, S., Leong, S. X., Rajaonson, E. M., Thiede, L. et al. (2024). In silico chemical experiments in the age of AI: From quantum chemistry to machine learning and back. Advanced Materials, 36(30), 2402369.

Apnavicius, J., Feng, Y., Flores, Y., Hassan, M., and McGuigan, M. (2021). Morse Potential on a Quantum Computer for Molecules and Supersymmetric Quantum Mechanics. arXiv:2102.05102v1 [quant-ph].

Huang, B., & Von Lilienfeld, O. A. (2021). Ab initio machine learning in chemical compound space. Chemical reviews, 121(16), 10001-10036.

Fedik, N., Nebgen, B., Lubbers, N., Barros, K., Kulichenko, M., Li, Y. W., et al. (2023). Synergy of semiempirical models and machine learning in computational chemistry. The Journal of chemical physics, 159(11).

Musil, F., Grisafi, A., Bartók, A. P., Ortner, C., Csányi, G., & Ceriotti, M. (2021). Physics-inspired structural representations for molecules and materials. Chemical Reviews, 121(16), 9759-9815.

Bradley, W., Kim, J., Kilwein, Z., Blakely, L., Eydenberg, M., Jalvin, J., et al. (2022). Perspectives on the integration between first-principles and data-driven modeling. Computers & Chemical Engineering, 166, 107898.

Orupattur, N. V., Mushrif, S. H., & Prasad, V. (2020). Catalytic materials and chemistry development using a synergistic combination of machine learning and ab initio methods. Computational Materials Science, 174, 109474.

Gyuk, N.J., Abubakar, M.S., and Abdu, S.G (2019). Using Asymptotic Iteration Method (AIM) for solving Differential Equations: Case Study of Vibrational States of Diatomic Molecule. Science World Journal. 14(3): 77-82

Majadifar, H., and Islampour ,R.(2012). Calculation of Franck-Condon Factors for a Number of Band Systems of Diatomic Molecules Using Hua Potential. Iran Journal of Chemistry and Chemical Engineering. 31(2):25-35

Fransson, E., Eriksson, F., & Erhart, P. (2019). Efficient construction and applications of higher-order force constant models. arXiv preprint arXiv:1902.01271.

Du, K. L., Jiang, B., Lu, J., Hua, J., & Swamy, M. N. S. (2024). Exploring Kernel Machines and Support Vector Machines: Principles, Techniques, and Future Directions. Mathematics (2227-7390), 12(24).

Mohammed, S. F., & Mahdi, G. J. M. (2024, March). Non-linear support vector machine classification models using kernel tricks with applications. In AIP Conference Proceedings (Vol. 3036, No. 1). AIP Publishing.

Fafalios, S., Charonyktakis, P., & Tsamardinos, I. (2020). Gradient boosting trees. Gnosis Data Analysis PC, 1.

Chintala, M. M. T. A.S. (2019). Optimizing predictive accuracy with gradient boosted trees infinancial forecasting. Turkish Journal of Computer and Mathematics Education (TURCOMAT), 10.

Akpojotor, G., Ehwerhemuepha , L., Echenim, M., and Akpojotor, F.(2010) Modeling and Visualization of Some Physics Phenomena with Vpython Simulations. African Journal of Physics.3:94-118

Akpojotor, G. and Ehwerhemuepha, L.(2012). An Overview of the Python African Computational Science and Engineering Tour Project.Invited Talk at PyCon US 2012 held at Silicon Valley-Santa Clara, California (USA) from March 7th-15th ,2012.Retrieved From. http://pyvideo.org/video/738/25-an-overview-of-the-python-african-computation

Sturm, E. J., Carbone, M. R., Lu, D., Weichselbaum, A., & Konik, R. M. (2021). Predicting impurity spectral functions using machine learning. Physical Review B, 103(24), 245118.

Jamali, B., Haghighat, E., Ignjatovic, A., Leitão, J. P., & Deletic, A. (2021). Machine learning for accelerating 2D flood models: Potential and challenges. Hydrological Processes, 35(4), e14064.

Janet, J. P., & Kulik, H. J. (2017). Resolving transition metal chemical space: Feature selection for machine learning and structure–property relationships. The Journal of Physical Chemistry A, 121(46), 8939-8954.

Dral, P. O. (2020). Quantum chemistry in the age of machine learning. The journal of physical chemistry letters, 11(6), 2336-2347.

Zhang, J., Cheng, P., Li, Z., Wu, H., An, W., & Zhou, J. (2023). A feature engineering method for machine learning inspired by quantum mechanics. In 2023 International Joint Conference on Neural Networks (IJCNN) (pp. 1-8). IEEE.

Nguyen, H. (2019). Support vector regression approach with different kernel functions for predicting blast-induced ground vibration: a case study in an open-pit coal mine of Vietnam. SN Applied Sciences, 1(4), 283.

Tawfik, S. A., Isayev, O., Spencer, M. J., & Winkler, D. A. (2020). Predicting thermal properties of crystals using machine learning. Advanced Theory and Simulations, 3(2), 1900208.

Zhang, Z., Zhao, Y., Canes, A., Steinberg, D., & Lyashevska, O. (2019). Predictive analytics with gradient boosting in clinical medicine. Annals of translational medicine, 7(7), 152.

Han, Y., Ali, I., Wang, Z., Cai, J., Wu, S., Tang, J., et al. (2021). Machine learning accelerates quantum mechanics predictions of molecular crystals. Physics Reports, 934, 1-71.

Tao, Y., Zeng, X., Fan, Y., Liu, J., Li, Z., & Yang, J. (2022). Exploring accurate potential energy surfaces via integrating Variational Quantum Eigensolver with machine learning. The Journal of Physical Chemistry Letters, 13(28), 6420-6426.