*JOURNAL OF NANO- AND ELECTRONIC PHYSICS*
Vol. **16** No 5, 05011(5pp) (2024)

*ЖУРНАЛ НАНО- ТА ЕЛЕКТРОННОЇ ФІЗИКИ*
Том **16** № 5, 05011(5cc) (2024)

OPEN ACCESS

## REGULAR ARTICLE

# An Innovative Classification Approach for Predicting Physical Properties in Nanoparticles

V.M. Tidake[1], P.M. Patare[2], P.B. Khatkale[3], A.A. Khatri[4,*], P.M. Yawalkar[5], S.S. Ingle[2], N.K. Darwante[6]

[1] *Department of MBA, Sanjivani College of Engineering, Kopargaon, SPPU, Pune, India*
[2] *Department of Mechanical Engineering, Sanjivani College of Engineering Kopargaon, SPPU, Pune, India*
[3] *Sanjivani University, Kopargaon, MH, India*
[4] *Department of Computer Engineering, Jaihind College of Engineering, Kuran, SPPU, Pune, MH, India*
[5] *Department of Computer Engineering, MET's Institute of Engineering, Nashik, India*
[6] *Department of Electronics and Computer Engineering, Sanjivani College Engineering, Kopargaon, SPPU, Pune, India*

(Received 02 June 2024; revised manuscript received 20 October 2024; published online 30 October 2024)

Zinc oxide (ZnO) nanoparticles (NP) are generating substantial attention across multiple areas due to the distinctive Structural and Molecular Features. Predicting and understanding these properties is crucial for designing effective applications in areas such as catalysis, sensors, and biomedical devices. Nanotechnology has emerged as a pivotal field, particularly in materials science, where the unique properties of NP are harnessed for various applications. Understanding and predicting the physical properties of NP, such as those in ZnO, is crucial for optimizing their performance. For the classification approach, we introduced a novel method, Bat based Random Forest (B-RF) to enhance the accuracy and efficiency of predicting major physical properties of ZnO NP. In this research, we utilize a relevant dataset encompassing various physical properties of ZnO NP. The model is fine-tuned to achieve optimal performance. The proposed Random Forest-based classification approach demonstrates superior predictive performance compared to traditional methods. Our model attains high accuracy and reliability in predicting diverse physical properties of ZnO NP. By the end of the study, our suggested approach outperforms other methods in terms of Accuracy (92.8%), Sensitivity (90.8%), and Specificity (93.9%). This can contribute to improve the overall performance and functioning of the existing model in a better way.

**Keywords**: Zinc oxide, Nanotechnology, Physical properties, Bat based Random Forest (B-RF).

DOI: 10.21272/jnep.16(5).05011 PACS number: 77.84.Bw

## 1. INTRODUCTION

Nanotechnology, which involves the modification of material at the nanoscale, has emerged as an innovative field with extensive consequences across multiple industries. Scientists have an important challenge in understanding and predicting the physical characteristics of NP, which are particles with size measured in nanometers [1]. The physical characteristics of NP, including their size, shape, surface area, and composition, are crucial for determining the behavior and efficiency of these tiny particles in many applications, spanning from medicine to electronics. Precisely predicting these characteristics is crucial for fully harnessing the promise of nanotechnology. At the nanoscale, substances present unique and frequently unexpected features that differ from those found in bulk [2]. This phenomenon occurs a consequence of quantum effects and increased surface-to-volume ratios, leading to modified physical, chemical, and biological interactions [3]. Therefore, it is crucial to create dependable techniques for forecasting the physical characteristics of NP, which requires multidisciplinary cooperation and inventive strategies. The intricacy of nanoscale formations presents a fundamental barrier to forecasting their characteristics. NP has complex forms, crystal structures, and surface changes, unlike larger materials, which contribute to their different characteristics [4]. To tackle this intricate issue, scientists have turned to computational techniques, utilizing complex algorithms and simulation to acquire a deeper understanding of the behavior of particles at the atomic and molecular scales. Computational tools play an essential part in predicting characteristics including stability, reactivity, and thermal conductivity [5]. These provide a vital connection between experiment and theory. The utilization of simulations of molecular dynamics and quantum mechanical computations has been highly beneficial in clarifying the mysteries surrounding the behavior of NP. Through the process of modeling, researchers may mimic the dynamic reactions of NP to

* Correspondence e-mail: khatrianand@gmail.com





Cite this article as: V.M. Tidake et al., *J. Nano- Electron. Phys.* 16 No 5, 05011 (2024) https://doi.org/10.21272/jnep.16(5).05011

outside stimuli by examining the interactions among atoms and molecules in NP. This provides a virtual laboratory environment that allows for the exploration of many situations [6]. These simulations serve the purpose of predicting physical attributes, but providing guidance to experimentalists in the process of developing and producing NP with specific features. The attempt of forecasting physical characteristics in NP encompasses theoretical development but also methods from experiments. Advanced methodologies, such as in situ microscopy and spectroscopy to allow for the direct and immediate monitoring of NP dynamics in different environments [7]. The combination of this data from experiments with computer models creates a synergistic approach that enhances our understanding of NP characteristics. Predicting the material properties of zinc oxide particles involves multiple challenges and constraints [8]. The accurate estimation of these features is heavily dependent on the accessibility of substantial empirical data, which may be limited or difficult to acquire at the nanostructures. Additionally, the fundamental complexity of particle systems, such as size-dependent impacts, surface modifications, and quantum mechanical events, offers significant computing difficulties [9, 10].

The rest of this article is divided into the following sections: Section 2, Related Works; Section 3, Methodology; Section 4, Performance evaluation; and Section 5, Conclusion.

## 2. RELATED WORK

The study [11] presented a machine learning (ML) method that predicted NP antibacterial properties with promising findings (R2 = 0.78). A literature study of 60 publications had yielded essential physicochemical (p-chem.) parameters and experimental conditions for in vitro experiments. Due to the non-linear connection among variables that were input and output, RF predicted the antibacterial impact better than other models.

The paper [12] examined how the artificial neural network (ANN) predicted the Heat conduction titanium dioxide-aluminum oxide NPusing spatial Temperature and Concentration levels. The thermal conductivity was predicted through the application of "self-organizing map (SOM) and Back Propagation-Levenberq-Marquardt (BP-LM)" methods. These methods were considered excellent predictors of thermal conductivity due to their results.

The article [13] employed multilayer perceptron (MLP) and long short-term memory (LSTM) connectivity to forecast metal-organic structure adsorption of gases abilities via a hypothetical collection of 130,000 MOF structures with natural Gas and greenhouse Gas absorption information for multiple pressures. Mixed NP was used to create effective deep learning models, proving that different nanomaterials could be used for deep learning.

The study [14] employed Deep Neural Network (DNN) to estimate Methylene Blue (MB)dye disappearance under TiO22 NPs. Additionally, the suggested DNN model integrated complicated input-output circumstances for optimal outcome prediction. The variables that were entered TiO22 NPs, ethylene glycol, and reaction time, while the response was MB dye elimination %. The Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and Standard Deviation (SD) findings showed that the DNN algorithm had lower error than Multiple Linear Regression.

The study [15] presented two-stage architecture for a ML-driven high quantities microfluidic system that produced NP of silver with the specified absorbance spectrum. After 120 circumstances, the computational approach converged to the desired spectrum using a Gaussian process-based Bayesian optimization (BO) and a DNN. When there was a desire for greater regression accuracy within the target, the recommendation was to consider altering the acquisition function during the second phase of the structure.

## 3. METHODOLOFY

### 3.1 Data Arrangement for ML

The initialization of the ZnO Np structure, containing 272Atomic Units, is presented in Fig. 1. The ZnO characterization of the NP involved a super cell of the hexagon crystal form with dimensions of 30*30*30.Statistical features of the data set including ZnONP were examined in this study. For reasons of clarity, we include the graphs corresponding to the lowest and highest data sets, despite the fact that we utilize atom shapes generated at temperatures ranging from 0 K to 998 K (100,250,450, 800). Every data set consists of 272 atoms at every temperature level, with each one characterized by its three-dimensional geometric dimensions. There is no linear correlation between the continuously variables being input, and every variable that is continuous follows a normal distribution. The Zn and O atoms are in equal quantity in the information (an equal distribution of 50% Zn and 50% O). When addressing classification concerns, these data sets are referred as distributed data.

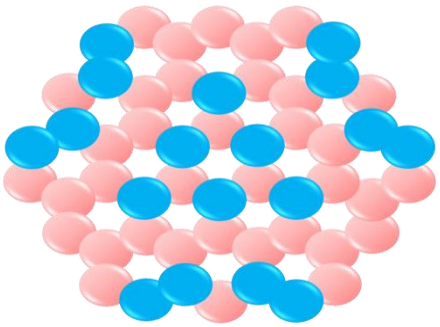

**Fig. 1** – The initial ZnO NP model structure comprises 272 atoms (pink is Oxygen, Blue is Zinc). (Source: Author)

### 3.2 Bat Algorithm

The Bat Algorithm (BA) has emerged as an innovative optimization technique inspiration from bat echolocation

behavior. The Bat Algorithm offers a robust optimization framework to enhance the accuracy of computational models. By leveraging the algorithm's ability to explore solution spaces efficiently, researchers can optimize the parameters of predictive models, leading to more accurate and reliable predictions of physical properties in ZnO NPs While echoes are used by most species of bats to locate their prey, not all bat varieties follow the exact same procedure. The microbat is a well-known example of an animal that uses echolocation. Consequently, the echolocation behavior is the first feature.

The second characteristic is the range at which the microbat searches for prey by sending out a set frequency,e_min, with a variable wavelength, $\lambda$, and loudness, $B_0$.

The volume may be changed in a variety of ways. For reasons of simplicity, it is assumed that the loudness may be adjusted between a positive, big B0 and $e_{min}$, the minimal constant value. Eqs. (1) – (3) simulate the motion of the virtual bat approach.

$$e_j = e_{min} + (e_{max} - e_{min}).\beta \quad (1)$$

$$u_j^s = u_j^{s-1} + (w_j^s - w_{best}).e_j \quad (2)$$

$$w_j^s = w_j^{s-1} + u_j^s \quad (3)$$

The suffixes min and max indicate the lowest and highest values, accordingly, where $e$ is the frequency that the bat uses to look for its prey. In the solution time, the $j-th$ bat's position is indicated by $w_j$. The bat's velocity, represented by$u_j$, is determined by $s$, which denotes the present version, $\beta$, a stochastic vector selected from an equal probability, $\beta \in [0,1]]$, and $w_{best}$, which shows the globally proximate optimal solution found thus extensively spread throughout the population as a whole. The formula $q_j q_{j\in[0,1]}$, where $j$ is the suffix indicating the $j^{th}$ bat. Every time, $q_j$ is contrasted with a randomly generated number. Randomised walk is a local search technique that is triggered when the random integer exceeds $q_j$. Equation (4) gives an unusual solution for the bat:

$$w_{new} = w_{old} + \varepsilon B^s \quad (4)$$

Where $B_j$ is the mean intensity of every bat at the current iteration, and $\varepsilon$ is a stochastic variable that falls between [-1, 1]. Only when the Earth-wide near-best solution is adjusted and the random generated value is less than $B_j$ are the audibility $B_j$ and pulse emissions rate $q_j$ changed following the update of bats' positions. Eqs. (5) and (6) act on the update of $B_j$ and $q_j$:

$$B_j^{s+1} = \propto . B_j^s \quad (5)$$

$$q_j^{s+1} = q_j^0[1 - f^{-\gamma s}] \quad (6)$$

### 3.3 Random Forests (RF)

RF is an ensemble learning that capitalizes on the collective power of several decision trees to enhance predictive accuracy and robustness. In the context of predicting physical properties in Zinc Oxide NP, RF excels in handling complex relationships and nonlinearities in the data, providing a comprehensive and accurate model.

At each stage of the building, one fully developed leaf is chosen. Each tree is built using half of the dataset information, randomly split in two. The structural features that determine the tree's form are accounted for when estimating divided sizes and divide attributes. The evaluation nodes suit the estimating techniques used in each tree leaf. Each tree has its assumptions about the data randomly split between the structure and estimate components. The Random forest algorithm is shown in Fig. 2.

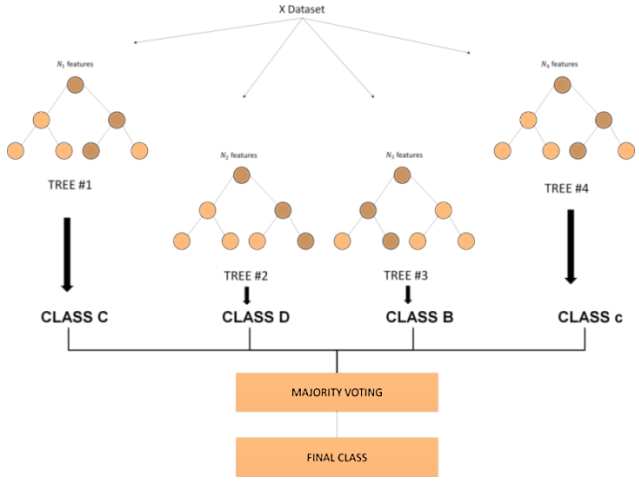


**Fig. 2** – Random forest model

The testing examples for the $i_1(y), i_2(y), .., i_l(y)$ classifiers were randomized from the distribution of the random vectors $Z, Y$. The wealth feature is presented as

$$Nh(Y,Z) = bw_l J(i_l(Y) = Z) - \frac{max}{k \neq Z} bh_l J(i_l(Y) = k) \quad (7)$$

Where the measured value is $J(.)$. The source of the mistake is

$$QF^* = Q_{Y,Z}(mg(Y,Z) < 0) \quad (8)$$

The probability over the $YZ$ dimension, indicated by where $Y, Z$ space $InRF, i_l(Y) = i(Y, \Theta_l)$

The margin feature for an RF is

$$mr(Y,Z) = Q_\Theta(i(Y,\Theta) = Z) - \frac{max}{k \neq Z} Q_\Theta(i(Y,\Theta)\} \quad (9)$$

Moreover, the set of classifiers $\{i(Y,\Theta)\}$ has a value of

$$T = F_{Y,Z} mr(Y,Z)\ K4.5 \quad (10)$$

### 3.4 Enhanced Prediction of Zinc Oxide Nanoparticle Physical Properties

The integration of the Bat algorithm and Random Forest algorithm presents a novel hybrid approach for predicting physical properties in Zinc Oxide NP. The Bat

algorithm, inspired by the echolocation behavior of bats, offers a powerful optimization technique to fine-tune the parameters of the Random Forest model, enhancing its predictive accuracy. By leveraging the strengths of both algorithms, this hybrid model aims to overcome challenges associated with traditional predictive methods for NP properties. Algorithm 1 shows the B-RF pseudocode.

***Algorithm 1:B-RF***

Initialize population of bats
Initialize Random Forest with parameters
Repeat until convergence:
    For each bat:
    Generate a new solution using Bat algorithm
    Evaluate the solution using Random Forest
    Update bat's position based on fitness
    Update Random Forest with new solutions
End loop

## 4. RESULT AND DISCUSSION

Receiver Operating Characteristic (ROC) curve is a Performance visualization for binary classification models, illustrating the balance among sensitivity and specificity at various decision points. For a comprehensive evaluation of algorithm performances, evaluating accuracy, specificity, and sensitivity. Fig. 3 depicts the proposed method outcome of ROC. Fig. 4 illustrates the outcome of Sensitivity-Specificity. Fig. 5 displays the Outcome of precision-recall.

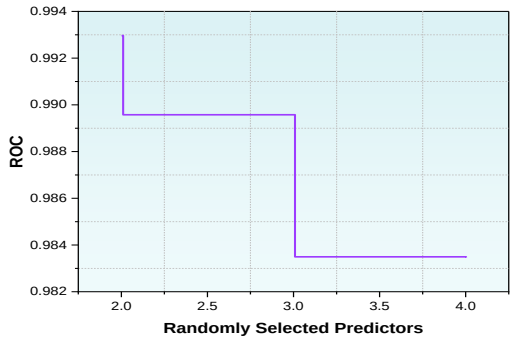


**Fig. 3** – Outcome of ROC

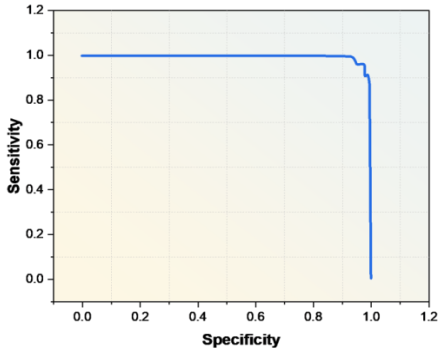


**Fig. 4** – Outcome of Sensitivity-Specificity

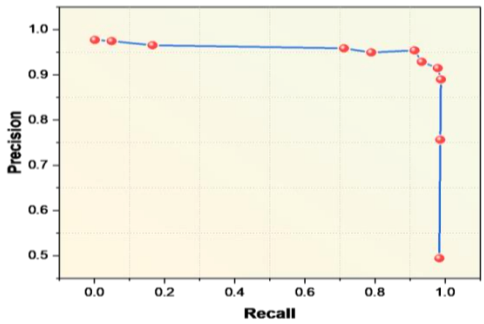


**Fig. 5** – Outcome of precision-recall

In this section, the performance evaluation of the proposed approach involves assessing it in terms of Accuracy, Sensitivity and Specificity, and conducting a comparative analysis with other existing methods, including "Flexible Discriminant Analysis (FDA) [21]and, Naive Bayes (NB), [21], K-Nearest Neighbor (KNN)[21]".

**Table 1** – Comparative evaluation

| Methods | Accuracy (%) | Sensitivity (%) | Specificity (%) |
|---|---|---|---|
| FDA | 88.9 | 82.6 | 77.4 |
| NB | 80.5 | 51.7 | 89.5 |
| KNN | 67.8 | 65.9 | 59.6 |
| B-RF [Proposed] | 92.8 | 90.8 | 93.9 |

Table 1 depict the comparative evaluation of accuracy. The entire accuracy of the classification or extraction procedure is measured by accuracy. When compared to currently existing methods such as FDA, NB and KNN, which have Accuracy values of 88.9%, 80.5% and 67.8%, respectively, The suggested B-RF achieves a higher accuracy of 92.8% for performs in estimating or forecasting the desired physical properties of the NP.

Specificity is a fundamental parameter utilized in the fields of statistics to evaluate predictive positive outcomes. A higher Specificity value implies that the model is making lower occurrences of false positives, indicating an accurate identification of positive cases.

Sensitivity is a fundamental metric used for evaluating the performance of predictive models, including those involved in predicting physical properties in Zinc Oxide NP.

## 5. CONCLUSION

In this study, we introduced a novel approach, Bat based Random Forest (B-RF), accurately estimated to physical properties in Zinc Oxide Nanoparticles. Experimental results showed Accuracy (92.8%), Sensitivity (90.8%), and Specificity (93.9%). The results of the proposed method were compared to the other utilized algorithms, and the outcomes of the evaluations showed that the suggested strategy was more effective for ability to predict key characteristics of zinc oxide nanoparticles. The quality of the data used for training and testing the model is crucial. Inaccuracies, noise,

or biases in the data may impact the model's predictive capabilities. In future research Continuously refine and optimize feature selection or engineering processes to identify the most relevant and influential features for predicting physical properties.

## REFERENCES


1. J.V. Staub de Melo, A.L. Manfro, B.S. Barra, N. Dell'Antonio Cadorin, W. Borba Broering, *Nanomaterials* **13** No 2, 288 (2023).
2. J. Vera, W. Herrera, E. Hermosilla, M. Díaz, J. Parada, A.B. Seabra, G. Tortella, H. Pesenti, G. Ciudad, O. Rubilar, *Antioxidants* **12** No 4, 784 (2023).
3. J. Jayaraman, K. Alagu, C.J. Raorane, P. Appavu, N. Joy, A. Mariadhas, J.A. Kumar, M. Rajasimman, *Fuel* **332**, 126170 (2023).
4. M. MuthuKathija, M.S.M. Badhusha, V. Rama, *Applied Surface Science Advances* **15**, 100400 (2023).
5. D.A. PD, D.R. Plashintania, R.M. Putri, I. Wibowo, Y. Ramli, S. Herdianto, A. Indarto, *Plos one* **18** No 7, e0289125 (2023).
6. W.A. Alhoqail, A.S. Alothaim, M. Suhail, D. Iqbal, M. Kamal, M.M. Asmari, A. Jamal, *Biomedicines* **11** No 2, 320 (2023).
7. M.A. Sani, A. Dabbagh-Moghaddam, G. Jahed-Khaniki, A. Ehsani, A. Sharifan, A. Khezerlou, M. Tavassoli, M. Maleki, *Journal of Food Measurement and Characterization*, 1 (2023).
8. Z. Efati, S.S. Shahangian, M. Darroudi, H. Amiri, S.I. Hashemy, M.R. Aghamaali, *Ceramics International* **49** No 20, 32568 (2023).
9. R. Huang, Z. Han, C. Ma, H. Liu, X. Huangfu, *Journal of Environmental Chemical Engineering* **11** No 1, 109069 (2023).
10. J.V. Staub de Melo, A.L. Manfro, B.S. Barra, N. Dell'Antonio Cadorin, W. Borba Broering, *Nanomaterials* **13** No 2, 288 (2023).
11. M. Mirzaei, I. Furxhi, F. Murphy, M. Mullins, *Nanomaterials* **11** No 7, 1774 (2021).
12. M. Sadeghzadeh, H. Maddah, M.H. Ahmadi, A. Khadang, M. Ghazvini, A. Mosavi, N. Nabipour, *Nanomaterials* **10** No 4, 697 (2020).
13. W. Guo, J. Liu, F. Dong, R. Chen, J. Das, W. Ge, X. Xu, H. Hong, *Nanomaterials* **12** No 19, 3376 (2022).
14. N. Amor, M.T. Noman, M. Petru, *Polymers* **13** No 18, 3104 (2021).
15. F. Mekki-Berrada, Z. Ren, T. Huang, W.K. Wong, F. Zheng, J. Xie, I.P.S. Tian, S. Jayavelu, Z. Mahfoud, D. Bash, K. Hippalgaonkar, *npj Computational Materials*, **7** No 1, 55 (2021).


## Інноваційний класифікаційний підхід для прогнозування фізичних властивостей наночастинок


V.M. Tidake[1], P.M. Patare[2], P.B. Khatkale[3], A.A. Khatri[4], P.M. Yawalkar[5], S.S. Ingle[2], N.K. Darwante[6]

[1] *Department of MBA, Sanjivani College of Engineering, Kopargaon, SPPU, Pune, India*
[2] *Department of Mechanical Engineering, Sanjivani College of Engineering Kopargaon, SPPU, Pune, India*
[3] *Sanjivani University, Kopargaon, MH, India*
[4] *Department of Computer Engineering, Jaihind College of Engineering, Kuran, SPPU, Pune, MH, India*
[5] *Department of Computer Engineering, MET's Institute of Engineering, Nashik, India*
[6] *Department of Electronics and Computer Engineering, Sanjivani College Engineering, Kopargaon, SPPU, Pune, India*



Наночастинки (NP) оксиду цинку (ZnO) привертають значну увагу в багатьох областях завдяки відмінним структурним і молекулярним особливостям. Прогнозування та розуміння цих властивостей має вирішальне значення для розробки ефективних застосувань у таких сферах, як каталіз, датчики та біомедичні пристрої. Нанотехнології стали ключовою сферою, особливо в матеріалознавстві, де унікальні властивості наночастинок використовуються для різних застосувань. Розуміння та прогнозування фізичних властивостей наночастинок, таких як ZnO, має вирішальне значення для оптимізації їх продуктивності. Для класифікаційного підходу ми представили новий метод, який підвищує точність і ефективність прогнозування основних фізичних властивостей ZnO NP. У цьому дослідженні ми використовуємо відповідний набір даних, що охоплює різні фізичні властивості наночастинок ZnO. Модель налаштована для досягнення оптимальної продуктивності. Запропонований підхід до класифікації демонструє кращу ефективність прогнозування порівняно з традиційними методами. Наша модель досягає високої точності та надійності в прогнозуванні різноманітних фізичних властивостей наночастинок ZnO. Запропонований підхід перевершує інші методи з точки зору точності (92,8 %), чутливості (90,8 %) і специфічності (93,9 %). Це може сприяти покращенню загальної продуктивності та функціонування існуючої моделі кращим чином.



**Ключові слова**: Оксид цинку, Нанотехнології, Фізичні властивості.