JOURNAL OF NANO- AND ELECTRONIC PHYSICS
Vol. **17** No 3, 03024(6pp) (2025)

ЖУРНАЛ НАНО- ТА ЕЛЕКТРОННОЇ ФІЗИКИ
Том **17** № 3, 03024(6cc) (2025)

OPEN ACCESS

REGULAR ARTICLE

Novel Model for Classifying the Toxicity of Metal Oxide Nanoparticles

Harshal P. Varade[1,*], Jaikumar M. Patil[2], Bhagyashree Ashok Tingare[3], P. William[4], Vaibhav D. Dabhade[5], R.A. Kapgate[6], Sachin K. Korde[7]

[1] *Department of Mechanical Engineering, Sanjivani College of Engineering, Kopargaon, MH, India*
[2] *Department of Computer Science and Engineering, Shri Sant Gajanan Maharaj College of Engineering, Shegaon, SGBAU, Amravati, India*
[3] *Department of Artificial Intelligence and Data Science, D Y Patil College of Engineering, Akurdi, Pune, India*
[4] *Department of Information Technology, Sanjivani College of Engineering, Kopargaon, MH, India*
[5] *MET Institute of Engineering, Nashik, India*
[6] *Department of Mechatronics Engineering, Sanjivani College of Engineering, Kopargaon, MH, India*
[7] *Department of Information Technology, Pravara Rural Engineering College, SPPU, Pune, India*

(Received 10 April 2025; revised manuscript received 21 June 2025; published online 27 June 2025)

Metal oxide nanoparticles (MeO_xNP) are receiving increasing attention in the last few years due to their various applications in electronics, medicine, and environmental remediation. However, their potential toxicity poses significant hurdles for safe usage. Therefore, this paper aims at developing a new artificial intelligence (AI)-based model for the efficient classification of the toxicity of MeO_xNP using a Dynamic Pelican Optimizer finetuned Random Forest (DPO-RF) technique. A database has been prepared considering different types of nanoparticles (NPs) such as Al_2O_3, CuO, Fe_2O_3, TiO_2, and ZnO, and the most important key physicochemical attributes. This model is followed by pre-processing using handling of missing values with imputation and performing standardization by applying the Z-score normalization. Features were extracted with principal component analysis (PCA) reducing dimension while keeping the vital information associated with toxicity in this model. The applied DPO-RF based model enhanced the feature selection of this model while achieving enhanced accuracy through adaptive exploration of this model. The results reflect the valid classification of MeO_xNP either as toxic or non-toxic, which implies a total accuracy of about 98.2 % for classes of toxicity, and a corresponding accuracy rate of about 98.5 % for classes of nontoxicity, which is offering some important implications for the assessment of potential risks while using the respective nanotechnology application.

Keywords: Metal oxide, Nanoparticles (NPs), Toxicity, Non-toxicity, Machine learning (ML).

DOI: 10.21272/jnep.17(3).03024 PACS numbers: 07.05.Mh, 77.84.Bw

1. INTRODUCTION

Nanoparticles (NPs) possess some unique, special physical and chemical characteristics. Some of their special features are characterized by tiny and quantum size influences. The NPs have wide applicability in many sectors, such as electronics, consumer goods (such as paints, textiles, and even cosmetics), personal hygiene products (such as sunscreen), and health-related issues within the framework of biomedicine. MeO_xNP make up the sizeable share of the overall market volume, and this is estimated to be at 80 percent. Indeed, lots of other nanomaterials (NMs) are in use (Na et al., 2020).

The scientific discipline of nanotechnology is concerned with the development and alteration of particles at a nanoscale, mainly in the range of 1-100 nanometers. It can be applied in making stronger, lighter, and more durable batteries, fuel cells, and solar cells compared to the traditional way. It can be used in computers, nanodiodes, and nano transistors (Nair et al., 2022; Zhang et al., 2022).

NPs are of a size smaller than 200 nm, which makes them suitable for biomedical applications. They also possess better colloidal stability and permeability through the blood-brain barrier. Their morphology needs to be characterized to utilize them effectively in diagnostics and therapeutics. Current applications of NPs include imaging molecular markers, genetic diseases, malignant tumors, photodynamic therapy, and drug delivery (Nikolova and Chavali 2020).

The traditional approaches regarding the determination of toxicity often follow experimental studies, whereby in vivo and in vitro testing are often solely used. Such in vitro and in vivo assessments usually face ethical and logistic limitations, especially when there is a variety of NMs presented and an array of conditions that

* Correspondence e-mail: varadeharshalmech@sanjivani.org.in

Cite this article as: H.P. Varade et al., *J. Nano- Electron. Phys.* 17 No 3, 03024 (2025) https://doi.org/10.21272/jnep.17(3).03024

might be encountered for the evaluation of toxicity. With ML and Deep learning (DL)-based models, such issues would be handled in the following way: large data sets would enable patterns to be discovered with toxicity levels predicted to show possible insights that guide the safe designs of MeO_xNP (Huang et al., 2022; Gakis et al., 2023).

2. RELATED WORKS

Chandrasekaran et al., (2024) advanced the protection of ZnO NPs by including factors together with Mg, Ca, Sr, and Ba. The produced NMs established reduced toxicity ranges each in vitro and in vivo, as well as possible antioxidant features. Because ZnCaO NMs consist of Ca_2 ions, they were much less dangerous to fish and mice, which increased material safety.

Zhou et al., (2023) used Deep learning (DL) capabilities and created an in silico version to forecast aquatic species' toxicity to metallic nanomaterials (MNMs). It was determined that variables, which include hydrodynamic diameter, primary length, light, and exposure time, all had an impact on ecotoxicity. More species-unique records were probably brought to the algorithm to decorate its prediction capabilities.

Zhang et al., (2023) anticipated the mixed toxicity of 7 steel-engineered NPs (ENPs) for Escherichia coli at numerous mixing ratios by using the use of literature and lab toxicity records. Two neural network community-quantitative structure-activity relationship (NN-QSAR) fashions and support vector machine (SVM)-QSAR models done nicely when ML strategies were used. For both inner and external datasets, the very best prediction capability changed into verified through a QSAR model based on NN and chemical descriptors.

Xiao created prediction models for the cellular toxicity of MeOxNP by automated ML, or autoML. AutoML generated models have better performance than ML models. Three different autoML systems' models performed satisfactorily; none of them outperformed the others. The performance of models constructed from higher-quality data was superior.

3. METHODOLOGY

This methodology outlines the comprehensive approach for evaluating the toxicity of metal oxide NPs through data sourcing, preprocessing, feature extraction, and classification. Key techniques such as principal component analysis (PCA) and the DPO-RF model are employed to enhance predictive accuracy and interpretability while effectively managing data complexities. Fig. 1 outlines the flow of methodology.

3.1 Data Set

The dataset, titled Toxicity Assessment MeO_xNP, was sourced from open source Kaggle (https://www.kaggle.com/datasets/apalania/toxicityassessment-meoxnp) and focuses on the computational hazard assessment of MeOxNP toxicity using ML techniques.

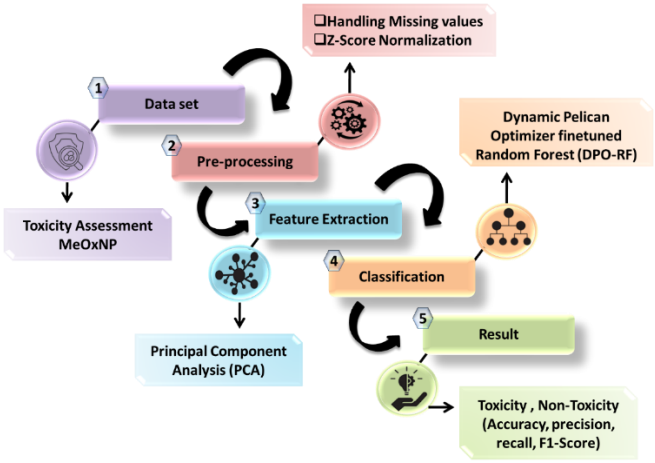

Fig. 1 – Methodology flow

It includes various NPs such as Al_2O_3, CuO, Fe_2O_3, TiO_2, and ZnO. The dataset encompasses more key features. These attributes provide a comprehensive overview of the chemical and physical properties of the NPs, along with their biological impact, enabling effective toxicity assessment and analysis through ML models.

3.2 Preprocessing

In predicting MeO_xNP toxicity, handling missing values and making use of Z-score normalization are crucial preprocessing steps. Imputation addresses statistics gaps, making sure dataset completeness, while Z-score normalization standardizes characteristic scales and mitigates outliers, making each descriptor comparably influential. Together, these techniques beautify statistics pleasant and model reliability, allowing accurate toxicity sample identification.

3.2.1 Handling Missing Values

In the preprocessing phase, addressing lacking values is essential to make certain the integrity and reliability of the dataset. Missing records can rise up from numerous assets, consisting of experimental mistakes, facts access troubles, or limitations in size strategies. To handle lacking values efficiently, numerous techniques can be hired. Imputation is a popular technique wherein missing values are substituted with approximated values derived from the records handy. Techniques, which include median, imply, or mode imputation, are used for mathematical capabilities, whilst uncompromising functions can be full of the majority of recurrent category.

3.2.2 Z-Score Normalization

The outlier problem is addressed by this statistical normalizing method. The feature values are transformed using the considered feature's mean and standard deviation. Specifically, values for the characteristic under consideration are converted using Equation (1) into new normalized values.

$$u' = \frac{u-\mu}{\sigma} \quad (1)$$

where σ and μ are the feature's standard deviation and the feature's mean value. When using the Z-score normalization approach, values that are precisely equal to the mean are translated to zero, values above the mean are shown as positive numbers, and values below the mean are shown as negative numbers.

3.3 Feature Extraction Using PCA

PCA is an essential technique for predicting the toxicity of MeOxNP. It transforms the authentic dataset into foremost additives that are uncorrelated variables ordered by way of variance. This reduces dimensionality at the same time as retaining good sized toxicity functions. PCA removes noise and redundant records, simplifying the dataset and permitting the version to recognition on applicable NP traits. It additionally aids in visualization of information in decreasing dimensions, decoding complex relationships amongst descriptors and their impact on toxicity effects.

3.4 Classification Using DPO-RF

The proposed DPO-RF model is appropriate for classifying the toxicity of MeOxNP. By combining DPO's strong seek capabilities with RF's excessive accuracy in handling complicated record styles, this model correctly selects and prioritizes relevant features whilst reducing the hazard of overfitting. DPO enhances RF's overall performance by refining feature choice through adaptive exploration and exploitation, supporting a strong and optimized type version. This technique aligns to gain dependable toxicity predictions, making sure each sturdy feature has relevance and excessive version interpretability in complex, multi-dimensional datasets.

3.4.1 Random Forest (RF)

RF is well-suited for predicting metal oxide NP toxicity due to its robustness in handling complex, high-dimensional datasets with diverse descriptors. By generating multiple uncorrelated decision trees, RF mitigates the risk of overfitting, ensuring stable predictions across varying samples. Its ensemble averaging technique reduces bias and variance, making it reliable for toxicity classification tasks. Furthermore, RF's inherent feature selection, by using random subsets, enhances model interpretability by identifying the most relevant descriptors, ultimately improving prediction accuracy and aligning well with the objective of reliable toxicity classification.

RF is an ensemble learning method utilized to generate predictive models. To increase prediction accuracy, ensemble techniques employ a variety of learning models. When using a RF, the method generates a whole forest of uncorrelated, random decision trees to find the optimal solution. Learning mistakes may often be explained by variation and bias. For instance, a large variance indicates that the model is only appropriate for a certain dataset (i.e., overfitting or instability), while a high bias leads to erroneous test findings. The training dataset $Y = \{y_1, \ldots . y_m\}$ is given labels$X = \{x_1, \ldots . x_m\}$. The training dataset is bagged repeatedly and randomly (K times) and then replaced with binary trees fitted to the samples. With$k = \{1, \cdots, K\}$, let y_k and x_k be the sampled dataset. Let S_a represent the binary tree that was trained about $y_k$and x_k. Two methods are available for making predictions on the test dataset, y, following training.

Calculating the average of each tree's predictions using (Equation 2).

$$\tilde{x} = \frac{1}{K}\sum_K S_a(\tilde{y}) \quad (2)$$

In the case of classification trees, obtaining the majority vote.

3.4.2 Dynamic Pelican Optimization (DPO)

DPO is a powerful tool for forecasting the toxicity of MO-NPs due to its ability to balance exploration and exploitation in complex search spaces. It uses a dynamic weighting mechanism to adjust search scales, enhancing solution quality and model convergence. This dual-phase approach allows DPO to efficiently locate global optima in optimization problems, making it an effective solution-finding tool.

Pelican Optimization

The PO is a novel stochastic optimization technique that draws inspiration from nature. It is renowned for its exceptional capacity to investigate and utilize the search space to find the global optimum. Recently, there has been a lot of interest in swarm-inspired algorithms, and the PO is particularly impacted by pelican behavior and foraging strategy. When hunting in the wild, pelicans frequently work together and take a multi-step method. Once they have identified where their prey is, they make a coordinated descent and then spread their wings. By forcing their prey to surface and migrate into shallower waters, the pelicans can catch their meal more easily.

Initialization

The PO is a population-based algorithm, initializing all pelican group members at random at the start of the optimization process using Equation (3). The population size, number of problematic variables, intervals, inferior and superior boundaries are represented by $W_{j,i}$.

$$W_{j,i} = KA_i + rand * (UB_i - LB_i) j = 1,2, \ldots M \quad (3)$$

Phase 1 (Exploration)

This phase involves pelicans searching for food sources, similar to pelicans hunting prey. The PO's ability to produce the prey's position at random improves its exploration skills. The objective function for the ith pelican candidate solution's new location is given in Equation (4), where O_i is the prey's location and $W_{j,i}^{O1}$ is the jth pelican's

most recent position. The J parameter greatly influences the PO's capacity to explore and methodically search the search space (Equation 5).

$$W_{j,i}^{O1} = \begin{cases} W_{j,i} + rand.\left(O_i - J.W_{j,i}\right), E_O < E_j; \\ W_{j,i} - rand.\left(O_i - J.W_{j,i}\right) else \end{cases} \quad (4)$$

$$W_j = \begin{cases} W_j^{O1}, E_j^{O1} < E_j; \\ W_j, \ else \end{cases} \quad (5)$$

Phase 2 (Exploitation)

Pelicans use their wings to drag fish upwards, collecting them in their neck pouch. This strategy increases fish catch in a designated area. The proposed PO improves convergence towards advantageous locations, increasing efficiency in local search and exploitation. The mathematical expression for this mimicking behavior is given in Equation (6), where s is the current iteration, s_{Max} is the maximum number of iterations, Q is a constant, and $W_{j,i}^{O2}$ is the jth pelican's most recent position.

$$W_{j,i}^{O2} = W_{j,i} + Q.\left(1 - \frac{s}{s_{Max}}\right).(2.rand - 1).W_{j,i} \quad (6)$$

The following Equation (7) then specifies how the result is modified in light of the new position.

$$W_j = \begin{cases} W_j^{O2}, E_j^{O2} < E_j; \\ W_j, else \end{cases} \quad (7)$$

Consequently, in accordance with Phase 2, $E_j^{O2}$is the value of this objective function, and W_j^{O2} is the jth pelican's most recent position.

3.5 System Configuration

The proposed method was implemented on a system with 16 GB of RAM and a 64-bit operating system to handle large datasets efficiently. Key libraries include NumPy for numerical operations, Pandas for data manipulation, Scikit-learn for ML, and Matplotlib for visualization, ensuring robust performance in forecasting the toxicity of metal oxide NPs.

3.6 Output Phase

This phase presents the classification performance metrics of the proposed method, highlighting its effectiveness in distinguishing between toxic and non-toxic metal oxide NPs. Table 1 summarizes the percentages of metrics for both target classes.

Table 1 – Performance of proposed classification

Target classes	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	Specificity (%)
Toxicity	98.2	97.8	98.2	98.3	97.5
Non-toxicity	98.5	98.6	98.5	98.4	98.9

4. RESULT AND DISCUSSION

Relevant performance metrics, including recall, precision, F1-score, accuracy, and specificity, are covered in this section, demonstrating the proposed method's effectiveness in classifying metal oxide NPs based on toxicity.

4.1 Accuracy

The accuracy (Equation 12) metrics for the proposed method indicate excellent performance in distinguishing between toxic and non-toxic classes of metal oxide NPs. The method achieved an accuracy of 98.2 % for the toxicity class and 98.5 % for the non-toxicity class. This high level of accuracy suggests that the model is effectively classifying instances, demonstrating its potential utility in real-world applications for toxicity assessment, where accurate predictions are crucial for safety and efficacy evaluations (Figure 2).

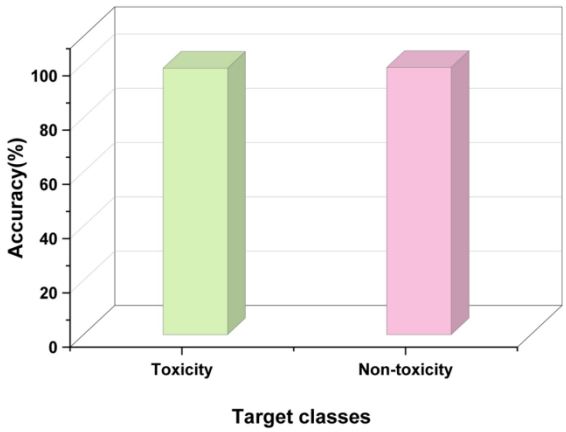

Fig. 2 – Accuracy performance of the proposed method

4.2 Precision and Recall

This parameter provides deeper insights into the performance. An accuracy of 97.8 % for the toxicity class indicates that the procedure is 97.8 % accurate for classifying toxicity. The recall signifies that the model classify 98.2 % of actual toxic instances. For non-toxicity, the precision is 98.6 %, meaning that almost all non-toxic predictions are accurate, while the recall is 98.5 %, indicating a high capability of detecting non-toxic instances. These metrics reflect the model's reliability in both minimizing false positives and maximizing true positives, which is crucial for ensuring that toxic NPs are correctly identified. Figure 3 (a) shows the performance of the proposed method with metric precision and Figure 3 (b) shows for recall.

4.3 F1-Score and Specificity

The F1-Score (Fig. 4(a)) is a harmonic mean of recall and precision, providing a single metric to assess the performance of the model on toxicity and non-toxicity classes.

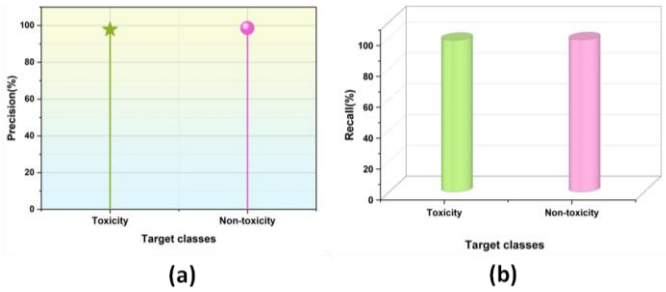

Fig. 3 – Performance of proposed method (a) precision and (b) recall

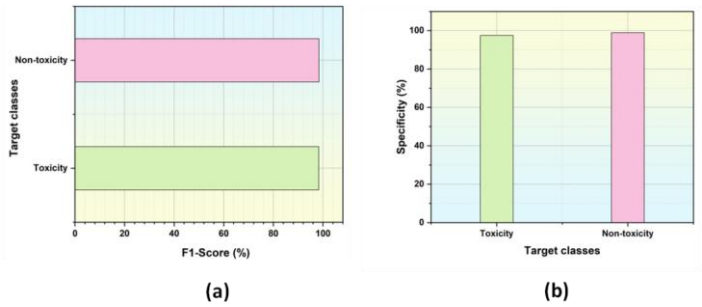

Fig. 4 – Performance of proposed method (a) F1-Score and (b) Specificity

The F1-Score of 98.3 % for toxicity and 98.4 % for non-toxicity indicates a well-balanced model with excellent overall performance. Specificity measures the proportion of true negatives accurately identified by the model. The specificity (Fig. 4(b)) for toxicity is 97.5 %, while for non-toxicity, it is 98.9 %. This represents that the model is very effective at identifying non-toxic NPs, with slightly lower performance for toxic ones, but indeed within acceptable ranges. Together, these metrics reinforce the robustness of the proposed method in accurately predicting toxicity while minimizing misclassifications.

5. CONCLUSION

This study successfully developed a novel AI-based model for classifying the toxicity of MeOxNP, utilizing the Dynamic Pelican Optimizer finetuned Random Forest (DPO-RF) approach. The model demonstrated high accuracy rates of 98.2 % for toxic and 98.5 % for non-toxic classes and interpretability, highlighting the importance of physicochemical properties in toxicity assessment. The findings contribute to enhanced risk assessment of NPs in various applications, supporting safer and more effective usage. Future research could expand the dataset to include additional NP types and explore alternative ML techniques to further improve classification performance and generalizability across different toxicological contexts. Limitations include the dataset's focus on specific metal oxide NPs and potential biases in the collected data. Future work should aim to incorporate a broader range of NPs and explore advanced ML techniques to enhance model robustness and applicability in diverse settings.

REFERENCES

1. M. Na, S.H. Nam, K. Moon, J. Kim, *Environ. Sci.: Nano* **10** No 1, 325 (2023).
2. G.M. Nair, T. Sajini, B. Mathew, *Talanta Open* **5**, 100080 (2022).
3. N. Zhang, G. Xiong, Z. Liu, *Front. Bioeng. Biotechnol.* **10**, 1001572 (2022).
4. M.P. Nikolova, M.S. Chavali, *Biomimetics* **5** No 2, 27 (2020).
5. T.Y. Suman, W.G. Li, D.S. Pei, *Toxicity of metal oxide nanoparticles* In *Nanotoxicity*, 107 (Elsevier: 2020).
6. H. Shi, Y. Pan, F. Yang, J. Cao, X. Tan, B. Yuan, J. Jiang, *Molecules* **26** No 8, 2188 (2021).
7. Y. Huang, X. Li, J. Cao, X. Wei, Y. Li, Z. Wang, X. Cai, R. Li, J. Chen, *Environ. Int.* **164**, 107258 (2022).
8. G.P. Gakis, I.G. Aviziotis, C.A. Charitidis, *Environ. Sci.: Nano* **10** No 3, 761 (2023).
9. K. Chandrasekaran, V. Kakani, V. Kokkarachedu, H.H.A. Syedahamed, S. Palani, S. Arumugam, A. Shanmugam, S. Kim, K. Kim, *Aquatic Toxicology* **267**, 106826 (2024).
10. Y. Zhou, Y. Wang, W. Peijnenburg, M.G. Vijver, S. Balraadjsing, W. Fan, *Environ. Sci. Technol.* **57** No 46, 17786 (2023).
11. F. Zhang, Z. Wang, W.J. Peijnenburg, M.G. Vijver, *Environ. Int.* **177**, 108025 (2023).
12. A.K. Halder, A. Melo, M.N.D. Cordeiro, *Chemosphere* **244**, 125489 (2020).
13. L. Sang, Y. Wang, C. Zong, P. Wang, H. Zhang, D. Guo, B. Yuan, Y. Pan, *Molecules* **27** No 18, 6125 (2022).
14. X. Xiao, T.X. Trinh, T.H. Yoon, *NANOTODAY-D-21-01586* (2022).

Нова модель для класифікації токсичності наночастинок оксидів металів

Harshal P. Varade[1], Jaikumar M. Patil[2], Bhagyashree Ashok Tingare[3], P. William[4], Vaibhav D. Dabhade[5], R.A. Kapgate[6], Sachin K. Korde[7]

[1] *Department of Mechanical Engineering, Sanjivani College of Engineering, Kopargaon, MH, India*
[2] *Department of Computer Science and Engineering, Shri Sant Gajanan Maharaj College of Engineering, Shegaon, SGBAU, Amravati, India*
[3] *Department of Artificial Intelligence and Data Science, D Y Patil College of Engineering, Akurdi, Pune, India*
[4] *Department of Information Technology, Sanjivani College of Engineering, Kopargaon, MH, India*
[5] *MET Institute of Engineering, Nashik, India*
[6] *Department of Mechatronics Engineering, Sanjivani College of Engineering, Kopargaon, MH, India*
[7] *Department of Information Technology, Pravara Rural Engineering College, SPPU, Pune, India*

Наночастинки оксидів металів (MeO_xNP) привертають дедалі більшу увагу в останні кілька років через їх різноманітне застосування в електроніці, медицині та відновленні навколишнього середовища. Однак їхня потенційна токсичність створює значні перешкоди для безпечного використання. Тому ця стаття спрямована на розробку нової моделі на основі штучного інтелекту (ШІ) для ефективної класифікації токсичності MeO_xNP з використанням методу точно налаштованого випадкового лісу (DPO-RF) динамічного оптимізатора Pelican. Була підготовлена база даних з урахуванням різних типів наночастинок (НЧ), таких як Al_2O_3, CuO, Fe_2O_3, TiO_2 та ZnO, а також найважливіших ключових фізико-хімічних характеристик. Ця модель супроводжується попередньою обробкою з використанням обробки відсутніх значень з імпутацією та стандартизацією шляхом застосування нормалізації Z-оцінки. Ознаки були вилучені за допомогою аналізу головних компонентів (PCA), зменшуючи розмірність, зберігаючи при цьому життєво важливу інформацію, пов'язану з токсичністю в цій моделі. Застосована модель на основі DPO-RF покращила вибір ознак цієї моделі, одночасно досягаючи підвищеної точності завдяки адаптивному дослідженню цієї моделі. Результати відображають дійсну класифікацію MeO_xNP як токсичної або нетоксичної, що передбачає загальну точність близько 98,2 % для класів токсичності та відповідний коефіцієнт точності близько 98,5 % для класів нетоксичності, що має деякі важливі наслідки для оцінки потенційних ризиків під час використання відповідного застосування нанотехнологій.

Ключові слова: Оксид металу, Наночастинки (НЧ), Токсичність, Нетоксичність, Машинне навчання (МН).