

Physicochemical Feature-Driven Nanotoxicity Prediction Using Supervised Machine Learning Algorithms.

Manuscript Info

Manuscript History

Received: xxxxxxxxxxxxxxxx
Final Accepted: xxxxxxxxxxxxxxxx
Published: xxxxxxxxxxxxxxxx

Key words:-

Nanotoxicity, Machine Learning,
Toxicity Prediction, Metal Oxide
Nanoparticles, Physicochemical
Properties.

Abstract

The widespread use of metal oxide nanoparticles across various industries has raised significant concerns regarding their potential toxicity. Conventional toxicological assessment methods remain time-intensive, costly, and limited in scalability. In this study, a machine learning-based framework was developed to classify nanoparticle toxicity using physicochemical descriptors. A dataset containing nine key features—such as dosage, surface area, and core size of the nanoparticles—was employed to train and evaluate six supervised learning algorithms: Decision Tree, Random Forest, Gradient Boosting, Logistic Regression, Support Vector Machine, and K-Nearest Neighbors. The analysis focused on five widely used metal oxide nanoparticles: Fe_2O_3 , TiO_2 , ZnO , CuO , and Al_2O_3 . The Decision Tree model achieved the highest classification accuracy (96.05%) and was noted for its interpretability and transparent decision rules. Model performance was assessed using ROC curves, precision-recall analysis, confusion matrices, and residual distributions, all of which confirmed the model's robustness and generalization capability. Feature importance analysis indicated that dosage, number of oxygen atoms, and electron affinity were the most significant predictors of toxicity. This approach enables accurate and interpretable nanotoxicity prediction and may serve as a valuable tool for risk-based assessment and the design of safer nanomaterials. The findings support the integration of data-driven methodologies into toxicological evaluation workflows.

Copy Right, IJAR, 2019,. All rights reserved.

1. Introduction:-

With the rapid expansion of nanotechnology, engineered nanoparticles have been increasingly employed in diverse sectors such as medicine [1], electronics [2], cosmetics [3], food packaging [4], and environmental remediation [5]. Among these, metal oxide nanoparticles—including ferric oxide (Fe_2O_3), titanium dioxide (TiO_2), zinc oxide (ZnO), copper oxide (CuO), and aluminum oxide (Al_2O_3) [6, 7]—have been widely adopted due to their favorable physicochemical properties such as high surface area [8], catalytic reactivity [9], chemical stability [10], and optical and electronic tunability [11, 12]. Despite their broad utility, growing concerns have been raised about the potential cytotoxic, genotoxic, and ecotoxic effects of these materials once released into biological and environmental systems [13–15].

Accurate and efficient toxicity assessment of nanoparticles is thus essential to ensure the safety of both human health and ecosystems [16]. However, conventional toxicity evaluation techniques, which typically involve in vitro or in vivo experiments, are time-consuming, expensive, and ethically challenging [17–20]. Moreover, the diverse

and complex interactions between nanoparticles and biological systems, driven by variations in size, shape, surface charge, solubility, and chemical composition, complicate the generalization of toxicological outcomes [21–24].

In response to these limitations, computational modeling approaches—particularly those based on machine learning—have emerged as powerful alternatives for predictive toxicology [25], and have been increasingly applied to predict a wide range of nanomaterial properties, including toxicity [26], band gap energy [27], solubility [28], bioavailability [29], catalytic activity [30], and zeta potential [31]. By learning from existing data, machine learning algorithms can uncover hidden patterns and associations between nanoparticle properties and toxicological outcomes, thereby enabling the development of predictive models that are both scalable and cost-effective [32, 33]. Such models have the potential to significantly reduce experimental burden while offering mechanistic insights into toxicity pathways [34].

In this study, the use of multiple supervised machine learning algorithms—including Decision Tree, Random Forest, Gradient Boosting, Logistic Regression, Support Vector Machine (SVM), and K-Nearest Neighbors—was investigated to classify the toxicity of five widely used metal oxide nanoparticles. Physicochemical descriptors of the nanoparticles were used as input features, and classification performance was evaluated through accuracy, error rate, and interpretability. Feature importance analyses were conducted to identify key predictors of toxicity. By integrating predictive modeling with interpretability, this work aimed to establish a reliable and transparent framework for nanotoxicity classification, ultimately contributing to the design of safer nanomaterials and the advancement of computational nanotoxicology.

2. Materials and methods:

The dataset used in this study was obtained from the Nanotoxicity Research Center, which provides experimentally validated nanotoxicity data for various engineered nanoparticles [35]. Specifically, it includes 875 samples corresponding to five widely used metal oxide nanoparticles: Fe_2O_3 , TiO_2 , ZnO , CuO , and Al_2O_3 . Each sample is described by nine physicochemical features: core size, hydrodynamic size, surface charge, surface area, electrical conductivity, exposure time, dosage, energy-related feature, and number of oxygen atoms. The classification target is binary, indicating whether a nanoparticle is toxic or non-toxic. The dataset contains approximately 470 toxic and 405 non-toxic entries, indicating a mild class imbalance.

2.1. Data Preprocessing

Prior to model training, non-informative columns such as nanoparticle identifiers were removed. The target labels were encoded numerically using label encoding (0 for non-toxic and 1 for toxic). All numeric features were standardized using z-score normalization via the StandardScaler from scikit-learn to ensure equal scaling and to improve model convergence. No missing values were present in the dataset.

2.2. Train–Test Splitting

The standardized dataset was divided into training and test sets using an 80/20 split. Stratified sampling was applied to maintain the relative distribution of toxic and non-toxic samples in both subsets. A fixed random seed (random_state=42) was used to ensure consistency and reproducibility across runs.

2.3. Machine Learning Models

A Random Forest classifier was employed as the primary predictive model for nanotoxicity classification due to its high performance and interpretability. The model was trained using the training set with default hyperparameters, including 100 estimators. Additionally, a Decision Tree estimator was extracted from the trained Random Forest

ensemble to provide an interpretable visualization of decision rules. All model development was performed using the scikit-learn library (v1.x) in Python 3.11.

2.4. Model Evaluation and Visualization

Model performance was assessed using several metrics: classification accuracy, error rate, precision, recall, and the area under the receiver operating characteristic curve (ROC-AUC). Additional evaluations included the confusion matrix, precision–recall curve, residual distribution, and a scatter plot of predicted versus actual probabilities to evaluate prediction confidence and calibration.

Feature importance was computed using the Gini impurity-based method native to the Random Forest algorithm. A decision tree from within the ensemble was visualized to explore how combinations of physicochemical features influenced toxicity predictions.

2.5. Implementation and Software Tools

All modeling, analysis, and visualizations were implemented in Python 3.11 using the following libraries: pandas, numpy, matplotlib, seaborn, and scikit-learn. Data visualizations—including class distribution, feature correlation heatmap, ROC and PR curves, residual histograms, and decision tree diagrams—were formatted using a journal-style configuration for readability and publication quality. Graphical elements were enhanced using serif fonts, higher DPI, and clear labeling conventions to align with common standards for scientific figures.

3. Results and discussion

3.1. Data Exploration

Prior to model training, the distribution of toxicity classes in the dataset was examined to identify any class imbalance that could impact model performance. As shown in Fig. 1, the dataset includes two target classes: "Toxic" and "nonToxic". Among the total samples, approximately 470 instances are labeled as Toxic, while around 405 instances are labeled as nonToxic. This indicates a mild imbalance, with toxic nanoparticles slightly outnumbering non-toxic ones.

Such imbalance, though not severe, may lead classification models to favor the majority class during training. Therefore, in addition to overall accuracy, we employed metrics such as precision, recall, ROC-AUC, and the precision-recall curve to ensure the selected model performs well on both classes. Addressing this class representation early allows for a more reliable evaluation of model behavior, especially in real-world scenarios where the cost of misclassifying a toxic nanoparticle may be significant [36, 37].

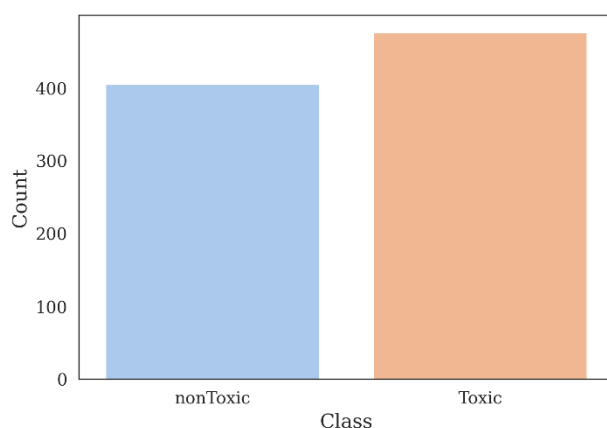


Fig. 1. Distribution of toxic and non-toxic class labels in the dataset, revealing a mild imbalance with toxic samples slightly outnumbering non-toxic ones

To investigate the relationships among the physicochemical descriptors of the nanoparticles, a Pearson correlation heatmap was generated. As illustrated in Fig. 2, most feature pairs show low to moderate correlation, indicating a minimal degree of multicollinearity. This is advantageous for machine learning applications, as it suggests that each feature contributes unique information to the classification task [38].

Among the strongest positive correlations observed, surface area and number of oxygen atoms exhibit a correlation coefficient of 0.52, implying that nanoparticles with larger surface areas tend to have more oxygen atoms — a trend that aligns with chemical intuition for metal oxides. Similarly, electron affinity shows moderate positive correlation with surface charge (0.51), and a notable negative correlation with number of oxygen atoms (-0.37), indicating underlying electronic or structural dependencies.

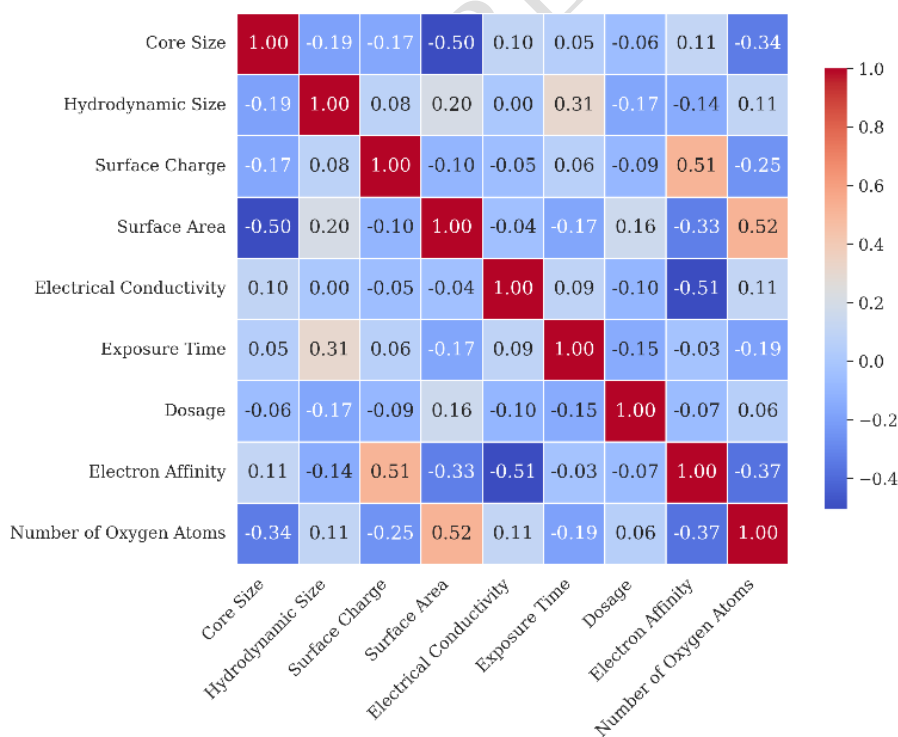


Fig. 2. Pearson correlation heatmap of physicochemical features, demonstrating low to moderate inter-feature correlation and minimal multicollinearity.

A moderate negative correlation of -0.50 was observed between core size and surface area, supporting the expected inverse relationship between particle size and surface-to-volume ratio. Other features such as electrical conductivity and exposure time displayed weak correlations with most variables, suggesting that they contribute independent variance and may serve as important predictors despite lacking strong pairwise correlations.

Overall, the heatmap confirms that the selected descriptors — including core size, hydrodynamic size, surface charge, surface area, electrical conductivity, exposure time, dosage, electron affinity, and number of oxygen atoms — are reasonably independent. This supports their inclusion in the model without significant concern for multicollinearity.

3.2. Model Evaluation

To evaluate the effectiveness of various machine learning algorithms in classifying nanotoxicity based on physicochemical features, six models were compared in terms of accuracy and error rate. The Decision Tree model achieved the highest classification accuracy at 96.05%, followed closely by Gradient Boosting (95.48%) and Random Forest (94.92%). The K-Nearest Neighbors algorithm also demonstrated strong performance with an accuracy of 94.92%, while SVM with RBF Kernel reached 89.27% accuracy. Logistic Regression, although less complex, achieved a respectable accuracy of 85.31%.

As shown in Fig. 3-a, Decision Tree and Gradient Boosting, outperformed the other classifiers in terms of accuracy. Ensemble models, known for reducing variance, consistently delivered higher predictive performance, suggesting their superior capability in modeling the complex relationships between physicochemical features and nanotoxicity [39]. Error rate analysis, shown in Fig. 3-b, mirrored the accuracy trend. The Decision Tree exhibited the lowest error rate (3.95%), indicating its robust generalization capability on this dataset. In contrast, Logistic Regression recorded the highest error (14.69%), likely due to its inability to capture non-linear feature interactions, which are essential in modeling complex biological phenomena such as nanotoxicity. For a clearer overview, Table. 1 summarizes the performance of all six models in terms of accuracy and error rate. Overall, ensemble methods such as Random Forest and Gradient Boosting demonstrated the most reliable performance, balancing high accuracy with low error. However, due to its highest predictive accuracy, low error rate, and interpretability, we decided to resume the remainder of the study using the Decision Tree model. Its ability to provide a transparent view of decision boundaries and feature contributions made it an ideal choice for further analysis, including feature importance assessment, decision rule extraction, and toxicological interpretation.

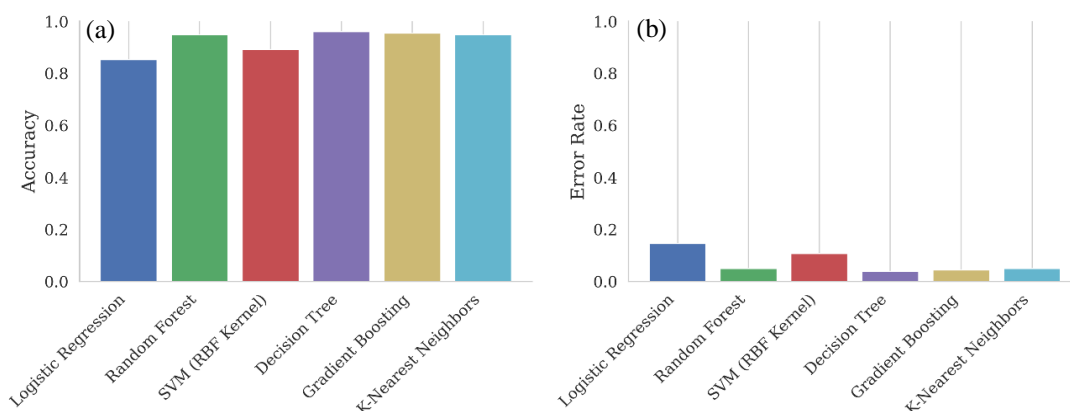


Fig. 3. Comparison of classification accuracy (a) and error rate (b) across six machine learning models applied to nanoparticle toxicity prediction, showing the Decision Tree with the highest accuracy (96.05%) and the lowest error, followed closely by Gradient Boosting.

Table 1. Accuracy and error rates of six machine learning models used for nanoparticle toxicity classification.

Model	Accuracy	Error rate
Logistic Regression	0.8531	0.1469
Random Forest	0.9492	0.0508
SVM (RBF Kernel)	0.8927	0.1073
Decision Tree	0.9605	0.0395
Gradient Boosting	0.9548	0.0452
K-Nearest Neighbors	0.9492	0.0508

To further assess the performance of the classification model, a confusion matrix was generated, as shown in Fig. 4. The matrix provides detailed insight into the model's prediction results by comparing actual class labels with predicted labels. Out of all test samples, 91 non-toxic nanoparticles were correctly identified, while only 2 were misclassified as toxic. Similarly, 77 toxic nanoparticles were correctly predicted, with 7 false negatives where toxic samples were incorrectly classified as non-toxic.

These results demonstrate that the model achieves a high true positive rate and a low false positive rate, consistent with the previously observed AUC of 0.98. Importantly, the relatively small number of false negatives is crucial in toxicological screening, where mislabeling a toxic substance as non-toxic could lead to hazardous consequences [40].

The confusion matrix thus highlights that the selected model (Decision Tree) maintains a strong balance between sensitivity (correctly identifying toxic samples) and specificity (correctly identifying non-toxic ones), making it well-suited for predictive nanotoxicology applications.

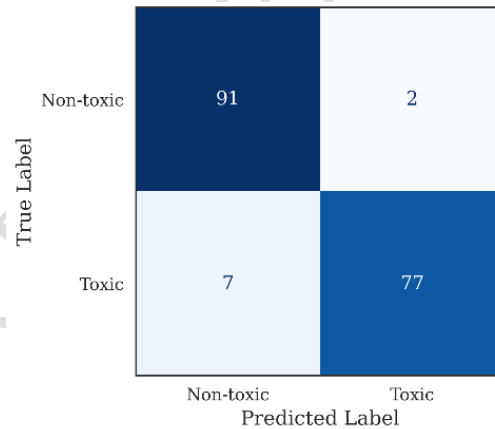


Fig. 4. Confusion matrix of the Decision Tree model, showing accurate classification of most samples and a low number of false predictions.

To evaluate the classification model's ability to distinguish between toxic and non-toxic nanoparticles, a Receiver Operating Characteristic (ROC) curve was plotted using the model's predicted probabilities on the test set. As shown in Fig. 5-a, the curve demonstrates a strong separation capability, with the true positive rate (TPR) rising sharply at very low false positive rates (FPR). The area under the curve (AUC) is 0.98, indicating excellent discriminatory power of the model.

An AUC close to 1.0 signifies that the model performs nearly perfectly in distinguishing between the two classes [41]. This high score confirms that the model correctly identifies toxic nanoparticles with minimal misclassification of non-toxic ones — a crucial consideration in nanotoxicology, where false negatives could lead to safety hazards in

biomedical or environmental applications. The ROC curve further supports the robustness of the selected features and the predictive strength of the Decision Tree model, justifying its use in the remainder of the analysis pipeline.

In addition to the ROC curve, a Precision-Recall (PR) curve was generated to further evaluate model performance, particularly under conditions of class imbalance. As shown in Fig. 5-b, the Decision Tree classifier maintained a consistently high precision across nearly the entire range of recall values. This indicates that the model was able to identify toxic nanoparticles with minimal false positives, even as it retrieved more true positives.

The near-horizontal plateau at precision ≈ 1.0 demonstrates that the model reliably classified most predicted toxic samples correctly, which is critical in applications where the cost of misidentifying a toxic nanoparticle as safe could be substantial. Only at very high recall levels does the precision drop slightly, reflecting the expected trade-off between sensitivity and precision in classification tasks [42]. This performance aligns with the confusion matrix and ROC analysis, confirming that the Decision Tree model not only performs well overall, but also remains highly dependable when prioritizing the accurate detection of true toxic cases.

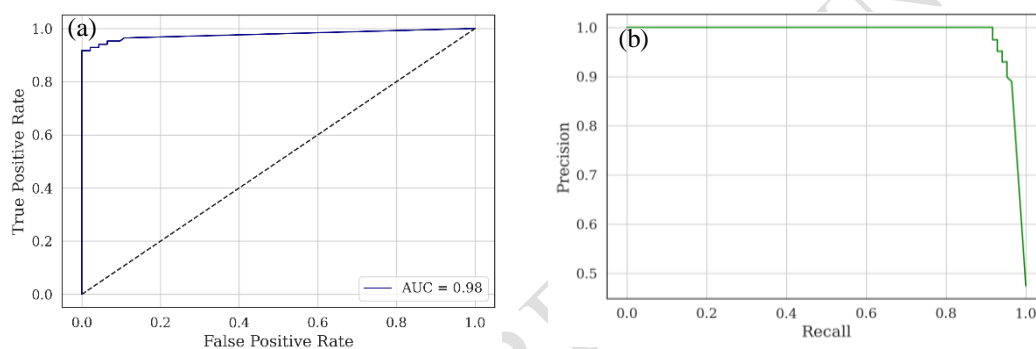


Fig. 5. (a) ROC and (b) Precision–Recall curves of the Decision Tree model, showing strong discriminatory power (AUC = 0.98) and high precision across recall levels.

To evaluate the calibration and confidence of the model’s probabilistic predictions, a residual analysis was performed. Residuals were calculated as the difference between the actual class label and the predicted probability output by the Decision Tree model. The resulting distribution is shown in Fig. 6-a. The plot reveals a strong peak centered around zero, indicating that the majority of predictions closely matched the actual outcomes. This narrow and sharply peaked distribution reflects minimal deviation between predicted probabilities and true labels, confirming that the model is well-calibrated [43]. Only a small number of residuals fall outside the -0.2 to $+0.2$ range, further highlighting the model’s tendency to make confident and accurate predictions.

Outliers on the far ends of the residual axis (near -1 or $+1$) represent rare cases where the model predicted a high probability for the incorrect class (e.g., predicting high toxicity for a non-toxic nanoparticle). The low frequency of such instances supports the overall robustness and reliability of the model’s predictive behavior.

To further evaluate the Decision Tree model’s behavior, a scatter plot of actual versus predicted probabilities was generated for the test set, as shown in Fig. 6-b. The blue points represent the model’s predicted probability that a given nanoparticle is toxic, while the red horizontal lines correspond to the actual class labels: “non-toxic” (class 0) and “toxic” (class 1). The plot reveals that the model tends to produce high-confidence predictions, with most probability values clustered around either 0.0 (strongly non-toxic) or 1.0 (strongly toxic). This reflects a clear separation in prediction space and indicates that the model is highly confident in its classifications. Predictions that align with the actual labels fall directly on the corresponding red lines, while misclassifications appear as blue dots deviating from the correct class level, particularly in the intermediate probability range between 0.3 and 0.7.

This distinct probability separation suggests that the model not only classifies accurately but also provides well-calibrated probabilistic estimates of toxicity. Such reliable probability outputs are essential for risk-based decision-making in nanotoxicology, where uncertainty can be just as critical as the classification itself [44].

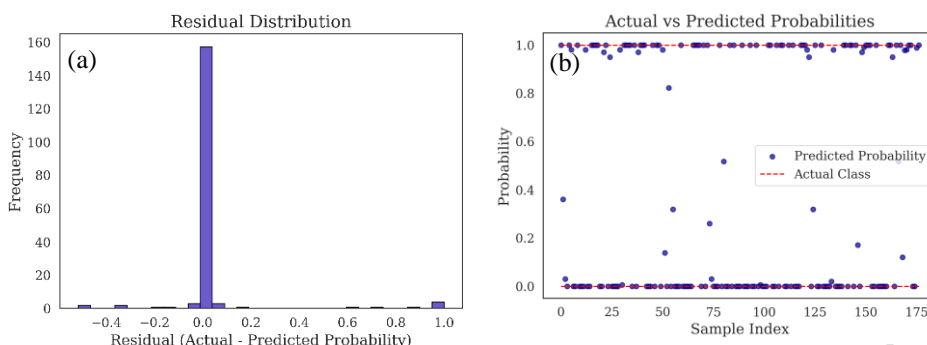


Fig. 6. (a) Residual distribution and (b) scatter plot of predicted vs. actual toxicity probabilities, highlighting the Decision Tree model's strong calibration and confident classification performance.

3.3. Model Interpretation

To interpret the contribution of each physicochemical feature to the Decision Tree model's predictions, feature importance scores were calculated based on the Gini impurity reduction criterion. As illustrated in Fig. 7, dosage was by far the most influential variable, contributing 39.3% of the total model importance. This finding underscores dosage as a critical determinant in nanoparticle toxicity, consistent with established toxicological principles where increased concentration correlates with greater biological impact [45].

Following dosage, the Number of Oxygen Atoms accounted for 15.6% of the predictive power, while electron affinity (e) contributed 13.8%. These features likely capture underlying chemical reactivity and molecular interactions relevant to cellular damage or oxidative stress. Exposure time and surface area followed with importance scores of 9.3% and 9.2%, respectively, emphasizing their moderate but meaningful roles in influencing toxic outcomes.

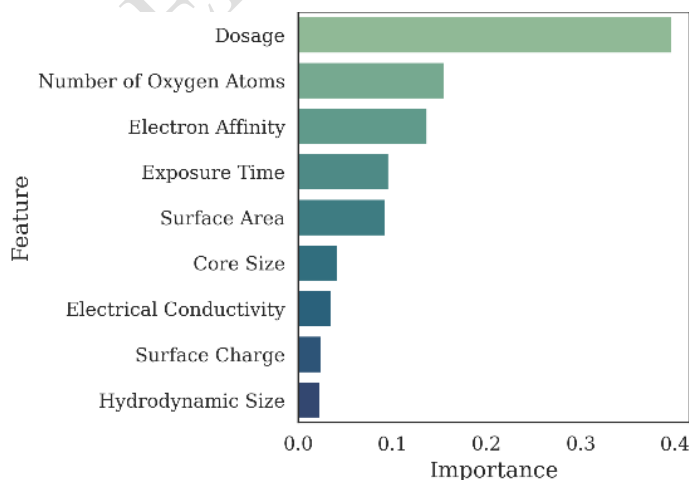


Fig. 7. Feature importance scores from the Decision Tree model, highlighting dosage, number of oxygen atoms, and electron affinity as the most influential predictors.

The remaining features had relatively minor influence: Core size (5.0%), electrical conductivity (4.3%), surface charge (3.0%), and hydrodynamic size (2.5%). These findings suggest that while all variables carry some predictive

weight, the model relies most heavily on dosage, oxygen content, and electronic properties to distinguish between toxic and non-toxic nanoparticles.

This ranking provides mechanistic insight into the physicochemical basis of nanotoxicity and confirms the utility of machine learning for interpretable and data-driven toxicological modeling.

To improve interpretability and examine how the model makes toxicity predictions, a representative decision tree was extracted from the trained Random Forest ensemble. As shown in Fig. 8, this tree reveals the sequential decision-making process the model uses to classify nanoparticles as toxic or non-toxic based on physicochemical features.

The root node begins with a split on surface area, where particles with values below -0.097 are more likely to be toxic. Subsequent branches incorporate key features such as exposure time, dosage, core size, and surface charge, which are used to progressively refine the classification. For example, particles with short exposure time and low dosage are typically classified as toxic, while combinations of low surface charge and small core size tend to result in a non-toxic classification.

Each node displays the Gini impurity, sample count, class distribution, and predicted class, providing insights into the decision purity at that point. Notably, many terminal nodes (leaves) reach high purity levels (Gini = 0), meaning all samples within those branches belong to the same class — a sign of strong separability in the data.

This visualization enhances the transparency of the machine learning approach, allowing researchers to trace specific decision paths and understand which combinations of features contribute most strongly to toxicity outcomes. In practice, such interpretable models can aid in forming hypotheses about nanoparticle behavior and guiding safer material design [46].

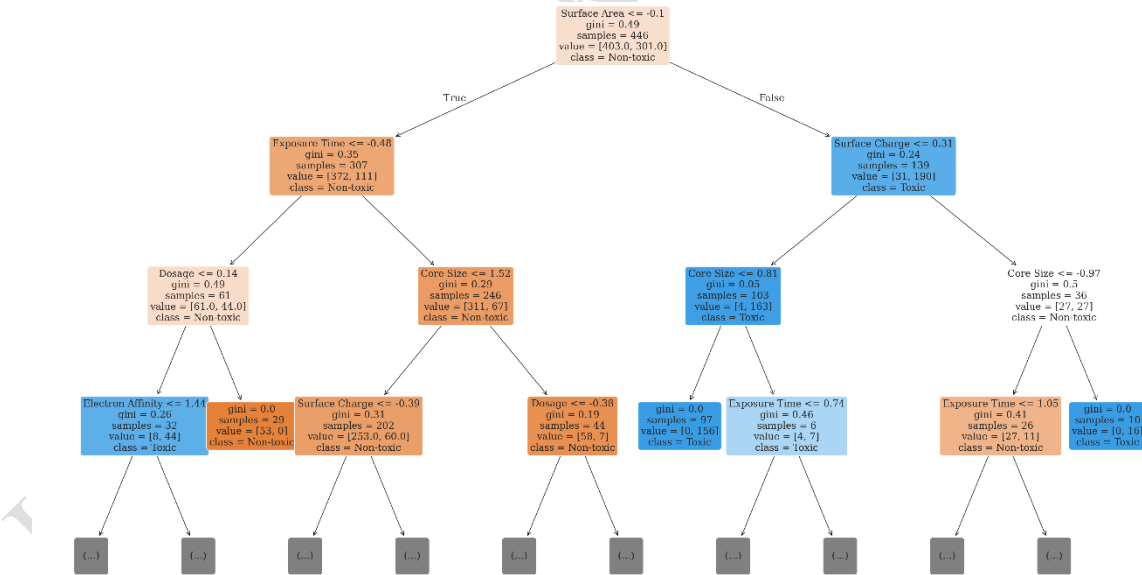


Fig. 8. Example Decision Tree extracted from the Random Forest model, illustrating decision rules based on physicochemical features for classifying nanoparticle toxicity.

4. Conclusion

This study demonstrated the utility of supervised machine learning models for the classification of metal oxide nanoparticle toxicity based on physicochemical properties. Among the six evaluated algorithms, the Decision Tree classifier achieved the highest accuracy (96.05%) and the lowest error rate (3.95%), while also offering strong interpretability. Through comprehensive model evaluation—including ROC and Precision-Recall analyses, confusion matrix interpretation, residual assessment, and decision rule visualization—the Decision Tree model proved to be both robust and transparent.

Feature importance analysis revealed that dosage, number of oxygen atoms, and electron affinity were the most influential descriptors in predicting nanotoxicity, highlighting the relevance of chemical and exposure-related parameters. Additionally, correlation analysis showed low multicollinearity among features, confirming their suitability for predictive modeling. By combining predictive accuracy with interpretability, this approach not only enables rapid toxicity screening but also provides mechanistic insight into the factors driving nanoparticle toxicity. These findings support the integration of machine learning into nanotoxicology workflows and contribute to the development of safer and more sustainable nanomaterials.

Acknowledgments

The author thanks the Kaggle website for providing the dataset and acknowledges the use of open-source Python libraries including scikit-learn, pandas, seaborn, and matplotlib.

Data and Code Availability

The dataset used in this study is publicly available from the UCI Machine Learning Repository via Kaggle at: <https://www.kaggle.com/datasets/ucimachinelearning/nanoparticle-toxicity-dataset>

All source code used for model development, evaluation, and visualization has been made openly available at the following GitHub repository: <https://github.com/PouyaPishkar/Toxicity.git>

Conflict of Interest

The authors declares that they have no conflicts of interest.

References

1. Yang X, Kubican SE, Yi Z, Tong S (2025) Advances in magnetic nanoparticles for molecular medicine. *Chem Commun* 3093–3108. <https://doi.org/10.1039/d4cc05167j>
2. He Z, Zhang Z, Bi S (2020) Nanoparticles for organic electronics applications. *Mater Res Express* 7:. <https://doi.org/10.1088/2053-1591/ab636f>
3. Cunha L, Monteiro J, Futuro A, et al (2025) Recycling PCBs for nanoparticles production with potential applications in cosmetics, cement manufacturing, and CO2 capture. *Waste Manag* 191:308–323. <https://doi.org/10.1016/j.wasman.2024.11.018>
4. Adeyemi JO, Fawole OA (2023) Metal-Based Nanoparticles in Food Packaging and Coating Technologies: A Review. *Biomolecules* 13:. <https://doi.org/10.3390/biom13071092>
5. Rafeeq H, Hussain A, Ambreen A, et al (2022) Functionalized nanoparticles and their environmental remediation potential: a review. *J Nanostructure Chem* 12:1007–1031. <https://doi.org/10.1007/s40097-021-00468-9>
6. Shawna L. Nations BS (2009) Acute and Developmental Toxicity of Metal Oxide Nanomaterials (ZnO, TiO₂, Fe₂O₃, and CuO) in *Xenopus laevis*. *Star*

- 263 7. Xiaojiang Yao, Fei Gao LD (2014) The application of incorporation model in γ -Al₂O₃-supported single and
264 dual metal oxide catalysts: A review. *Chinese J Catal* 35:108–119. <https://doi.org/10.1016/S1872>
- 265 8. Wei C, Sun S, Mandler D, et al (2019) Approaches for measuring the surface areas of metal oxide
266 electrocatalysts for determining their intrinsic electrocatalytic activity. *Chem Soc Rev* 48:2518–2534.
267 <https://doi.org/10.1039/c8cs00848e>
- 268 9. Yang D, Gates BC (2024) Characterization, Structure, and Reactivity of Hydroxyl Groups on Metal-Oxide
269 Cluster Nodes of Metal–Organic Frameworks: Structural Diversity and Keys to Reactivity and Catalysis.
270 *Adv Mater* 36:. <https://doi.org/10.1002/adma.202305611>
- 271 10. Gunjekar JL, Kim TW, Kim HN, et al Mesoporous Layer-by-Layer Ordered Nanohybrids of Layered
272 Double Hydroxide and Layered Metal Oxide: Highly Active Visible Light Photocatalysts with Improved
273 Chemical Stability. 1–9
- 274 11. Roman BJ, Shubert-Zuleta SA, Milliron DJ (2024) Tunable optical response of plasmonic metal oxide
275 nanocrystals. *MRS Bull* 49:1032–1044. <https://doi.org/10.1557/s43577-024-00785-8>
- 276 12. Chang BS, Martin A, Thomas B, et al (2020) Synthesis of Interface-Driven Tunable Bandgap Metal Oxides.
277 *ACS Mater Lett* 2:1211–1217. <https://doi.org/10.1021/acsmaterialslett.0c00251>
- 278 13. Simeone FC, Costa AL (2019) Assessment of cytotoxicity of metal oxide nanoparticles on the basis of
279 fundamental physical-chemical parameters: A robust approach to grouping. *Environ Sci Nano* 6:3102–3112.
280 <https://doi.org/10.1039/c9en00785g>
- 281 14. Nazanin Golbamaki, Bakhtiyor Rasulev, Antonio Cassano, Richard L. Marchese Robinson, Emilio
282 Benfenati, Jerzy Leszczynski MTDC (2015) Mini-Review Genotoxicity of metal oxide nanomaterials:
283 Review of Recent data and discussion of possible mechanisms. *Nanoscale*
- 284 15. Renzi M, Blašković A (2019) Ecotoxicity of nano-metal oxides: A case study on daphnia magna.
285 *Ecotoxicology* 28:878–889. <https://doi.org/10.1007/s10646-019-02085-3>
- 286 16. Medina-Ramirez IE, Jimenez-Chavez A, De Vizcaya-Ruiz A (2022) Toxicity of nanoparticles. *Antimicrob*
287 *Act Nanoparticles Appl Wound Heal Infect Treat* 249–284. [https://doi.org/10.1016/B978-0-12-821637-](https://doi.org/10.1016/B978-0-12-821637-8.00006-7)
288 [8.00006-7](https://doi.org/10.1016/B978-0-12-821637-8.00006-7)
- 289 17. Neeraja Revi ODO and DB (2023) In Vitro, In Vivo and Ex Vivo Models for Toxicity Evaluation of
290 Nanoparticles: Advantages and Disadvantages. *Toxic Nanoparticles-Recent Adv New Perspect* 13
- 291 18. Bakand S, Winder C, Khalil C, Hayes A (2005) Toxicity assessment of industrial chemicals and airborne
292 contaminants: Transition from in vivo to in vitro test methods: A review. *Inhal Toxicol* 17:775–787.
293 <https://doi.org/10.1080/08958370500225240>
- 294 19. ANDRE NEL, TIAN XIA, HUAN MENG, XIANG WANG, SIJIE LIN, ZHAOXIA JI and HZ (2013)
295 Nanomaterial Toxicity Testing in the 21st Century: Use of a Predictive Toxicological Approach and High
296 Throughput Screening ANDRE. *Acc Chem Res*. <https://doi.org/10.1021/ar300022h>.Nanomaterial
- 297 20. Ukelis U, Kramer PJ, Olejniczak K, Mueller SO (2008) Replacement of in vivo acute oral toxicity studies
298 by in vitro cytotoxicity methods: Opportunities, limits and regulatory status. *Regul Toxicol Pharmacol*
299 *51*:108–118. <https://doi.org/10.1016/j.yrtph.2008.02.002>
- 300 21. Sayes CM, Warheit DB (2009) Characterization of nanomaterials for toxicity assessment. *Wiley Interdiscip*
301 *Rev Nanomedicine Nanobiotechnology* 1:660–670. <https://doi.org/10.1002/wnan.58>
- 302 22. Fubini B, Ghiazza M, Fenoglio I (2010) Physico-chemical features of engineered nanoparticles relevant to
303 their toxicity. *Nanotoxicology* 4:347–363. <https://doi.org/10.3109/17435390.2010.509519>

23. Karakoti AS, Hench LL, Seal S (2006) The potential toxicity of nanomaterials - The role of surfaces. *Jom* 58:77–82. <https://doi.org/10.1007/s11837-006-0147-0>
24. Andrew D. Maynard DBW and MAP (2011) The New Toxicology of Sophisticated Materials: Nanotoxicology and Beyond. *Toxicol Sci* 1–37
25. Wu X, Zhou Q, Mu L, Hu X (2022) Machine learning in the identification, prediction and exploration of environmental toxicology: Challenges and perspectives. *J Hazard Mater* 438:. <https://doi.org/10.1016/j.jhazmat.2022.129487>
26. Lin Z, Chou WC (2022) Machine Learning and Artificial Intelligence in Toxicological Sciences. *Toxicol Sci* 189:7–19. <https://doi.org/10.1093/toxsci/kfac075>
27. Pishkar P Predicting Band Gap of Carbon and Nitrogen-Based Photocatalysts Using Machine Learning. 1–11
28. Liu B, Yu Y, Liu Z, et al (2025) Prediction of CO2 solubility in aqueous amine solutions using machine learning method. *Sep Purif Technol* 354:129306. <https://doi.org/10.1016/j.seppur.2024.129306>
29. Rezaei F, Rastegari Mehr M, Shakeri A, et al (2024) Predicting bioavailability of potentially toxic elements (PTEs) in sediment using various machine learning (ML) models: A case study in Mahabad Dam and River-Iran. *J Environ Manage* 366:121788. <https://doi.org/10.1016/j.jenvman.2024.121788>
30. Sinha P, Jyothirmai M V., Abraham BM, Singh JK (2024) Machine learning driven advancements in catalysis for predicting hydrogen evolution reaction activity. *Mater Chem Phys* 326:129805. <https://doi.org/10.1016/j.matchemphys.2024.129805>
31. Muneer R, Hashmet MR, Pourafshary P, Shakeel M (2023) Unlocking the Power of Artificial Intelligence: Accurate Zeta Potential Prediction Using Machine Learning. *Nanomaterials* 13:1–17. <https://doi.org/10.3390/nano13071209>
32. Mehendale P (2024) Scalable Architecture for Machine Learning Applications Available online www.jsaer.com Scalable Architecture for Machine Learning Applications. <https://doi.org/10.5281/zenodo.13753585>
33. Vajpayee A, Technologies M, Mohan R, et al (2024) BUILDING SCALABLE DATA. 15:308–320
34. Moukheiber L, Mangione W, Moukheiber M, et al (2022) Identifying Protein Features and Pathways Responsible for Toxicity Using Machine Learning and Tox21: Implications for Predictive Toxicology. *Molecules* 27:1–21. <https://doi.org/10.3390/molecules27093021>
35. Kaggle (2021) UCI Machine Learning Repository, “Nanoparticle Toxicity Dataset,.” <https://www.kaggle.com/datasets/ucimachinelearning/nanoparticle-toxicity-dataset>
36. Miao J, Zhu W (2022) Precision–recall curve (PRC) classification trees. *Evol Intell* 15:1545–1569. <https://doi.org/10.1007/s12065-021-00565-2>
37. Varghese A, Agyeman-Badu G, Cawley M (2020) Deep learning in automated text classification: a case study using toxicological abstracts. *Environ Syst Decis* 40:465–479. <https://doi.org/10.1007/s10669-020-09763-2>
38. BUYRUKOĞLU S, AKBAŞ A (2022) Machine Learning based Early Prediction of Type 2 Diabetes: A New Hybrid Feature Selection Approach using Correlation Matrix with Heatmap and SFS. *Balk J Electr Comput Eng* 10:110–117. <https://doi.org/10.17694/bajece.973129>
39. Xiao X, Trinh TX, Gerelkhuu Z, et al (2024) Automated machine learning in nanotoxicity assessment: A comparative study of predictive model performance. *Comput Struct Biotechnol J* 25:9–19. <https://doi.org/10.1016/j.csbj.2024.02.003>

- 346 40. Markoulidakis I, Markoulidakis G (2024) Probabilistic Confusion Matrix: A Novel Method for Machine
347 Learning Algorithm Generalized Performance Analysis. *Technologies* 12:.
348 <https://doi.org/10.3390/technologies12070113>
- 349 41. Fawcett T (2006) An introduction to ROC analysis. *Pattern Recognit Lett* 27:861–874.
350 <https://doi.org/10.1016/j.patrec.2005.10.010>
- 351 42. Saito T, Rehmsmeier M (2015) The precision-recall plot is more informative than the ROC plot when
352 evaluating binary classifiers on imbalanced datasets. *PLoS One* 10:1–21.
353 <https://doi.org/10.1371/journal.pone.0118432>
- 354 43. Niculescu-Mizil A, Caruana R (2005) Predicting good probabilities with supervised learning. *ICML 2005 -*
355 *Proc 22nd Int Conf Mach Learn* 625–632. <https://doi.org/10.1145/1102351.1102430>
- 356 44. Zhou Y, Wang Y, Peijnenburg W, et al (2024) Application of Machine Learning in Nanotoxicology: A
357 Critical Review and Perspective. *Environ Sci Technol*. <https://doi.org/10.1021/acs.est.4c03328>
- 358 45. Chen Q, Riviere JE, Lin Z (2022) Toxicokinetics, dose–response, and risk assessment of nanomaterials:
359 Methodology, challenges, and future perspectives. *Wiley Interdiscip Rev Nanomedicine Nanobiotechnology*
360 14:1–42. <https://doi.org/10.1002/wnan.1808>
- 361 46. De Felice F, Crocetti D, Parisi M, et al (2020) Decision tree algorithm in locally advanced rectal cancer: an
362 example of over-interpretation and misuse of a machine learning approach. *J Cancer Res Clin Oncol*
363 146:761–765. <https://doi.org/10.1007/s00432-019-03102-y>

UNDER PEER REVIEW