Analysis of properties leading to cytotoxicity of silver nanoparticles using Machine Learning

Neeru Bhagat[1]*, Anjana S Desai[1], Aparna Ashok[1], Mayur Gaikwad[1], Selvan Ravindran[1],Brajesh Pandey[1]

*[1]Symbiosis Institute of Technology, SymbiosisInternational (Deemed) University, 412115, Pune *neeru.bhagat@sitpune.edu.in (Neeru Bhagat)*

Abstract

The last decade has observed an exponential increase in the application of silver nanoparticles (Ag-NPs) synthesized by various routes. The extensive use of Ag-NPs has raised concerns about their potential impact on human health, which is debatable in terms of toxicity whenit comes to direct contact with the human body. The present study comprehends the relationship between the physical parameters of Ag-NPsand theircytotoxicity on normal and carcinoma cell lines. This study deals with a meta-analysis of the cytotoxicity data of Ag-NPs by collecting the heterogeneous input features from the literature using machine learning tools, decision tree (DT), and random forest (RF). Significance of the selected input features was ranked using the permutation-based measure known as mean decrease Gini (MDG). Performances of the models wereevaluated through regression metrics, root mean square error (RMSE) and square of standard deviation (R^2). Decision tree and random forest assisted in classifying the input parameters by asserting certain threshold values.The obtained high value of R^2 (DT-0.98, RF-0.88) and low value of RMSE (DT-3.9, RF-8.9) confirmed the accurate predictionand classification of toxicity and non-toxicity on normal and carcinoma cell lines. An attempt has been made to understand the toxicity of Ag-NPs to normal and carcinoma cell lines.

Keywords. Ag-NPs, Cytotoxicity, HEK-293, PC 12 cell line, Machine learning, Decision tree, Random Forest, K-means clustering, Regression metrics

1. INTRODUCTION

Ag-NPs are widely employed in consumer goods, including cosmetics and ointments, and are in great demand [1]. Ag-NPs have a wide range of applications in medicine, pharmaceuticals, pharmacology, biotechnology, electronics, engineering, energy, magnetic fields, and environmental cleanup [2]. In addition to these uses, Ag-NPs have become more significant in the industrial field of textiles, food, consumer goods, etc., due to their inborn and powerful antibacterial capabilities [3]. Ag-NPs are currently being investigated for medical devices, feminine hygiene products, paints, sunscreen, biosensors, textiles, and electronics [1-3]. Despite these advantages, Ag-NPs have also been found to be toxic.

Because of their potent oxidative properties, silver ions (Ag+) are released, which have detrimental effects on biological systems in the form of cytotoxicity, genotoxicity, immunological reactions, and even cell death (apoptosis) [4-7].The precise methods by which NPs interact with biological entities are yet unknown, and the mechanism of Ag-NPs cytotoxicity is not fully understood [8-11]. Recent studies have shown that the physical properties of Ag-NPs, such as their particle size, NPs dosage, and agglomeration, are crucial in determining their cytotoxicity. Investigations revealed that the kinetic development of NPs at each stage of synthesis greatly influences their physical properties [12,13]. To fine-tune the cytotoxicity behaviour of Ag-NPs, it is essential to determine the predictive correlation between cytotoxicity and the physical characteristics that arises due to the synthesis techniques. Hence, we report a comprehensive analysis for predicting and classifying cytotoxicity associated with Ag-NPs. This has been achieved by two machine learning algorithms, (i) decision tree (DT) and (ii) random forest (RF), run by datamining found in the literature. The findings are expected to reveal the key parameters that affect the cytotoxicity of Ag-NPs. Several nanoparticles are considered in the data mining and may affect the predictive results of machine learning as different nanoparticle exhibits various physical properties. The present work is inspired by our previous work on Ag-NPs synthesized using turmeric extract [14]. The input parameters are chosen based on Ag-NPs on normal cell lines and carcinoma cell lines using supervised learning, i.e. Decision Tree (DT) and the Radom Forest (RF).

2. METHODOLOGY

The preferred keywords for collecting data were "silver nanoparticles for cytotoxicity". To avoid duplication, we limited our search to only two databases, ScienceDirect and MDPI. Additionally, we focused on the most recent cytotoxicity study and collected the data from October 2019 to April 2022. We were only concerned with current research articles. From ScienceDirect, we located about 38 research articles, and from MDPI, about 21. Around 40 of the 59 articles could be accessed, and we acquired 2000 datasets (cell viability) and their corresponding concentrations of NPs. Out of them, 485 datasets were excluded due to cytotoxicity assays other than the MTT assay. In the end, 200 data were eliminated since the concentrations were not measured in µg/mL but in other units. The result of data mining is a final yield of 1315 data points from 40 articles with 11 different input features, including reducing agents, particle size, zeta potential, cell line type (normal/carcinoma), hydrodynamic diameter, wavelength, morphology, exposure time and concentration of silver based NPs for assessing cytotoxicity using MTT assay. The plant and non-plant mediated extracts are utilized as biosynthesizing agents in producing green AgNPs that might affect toxicity. The phytochemicals present in such synthesis agents may interact with other factors to affect the toxic effects of AgNPs. This could be a possible input feature for further negotiating the different machine learning algorithms to enhance their performance metrics.

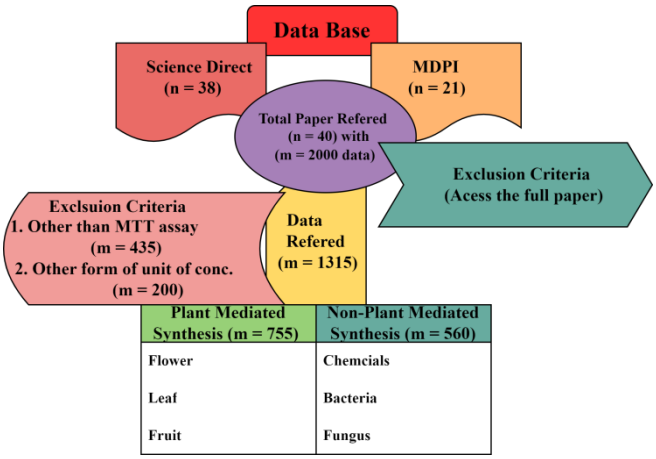


Figure 1. Block diagram representing the selection procedure of the research articles

3. RESULTS AND DISCUSSION

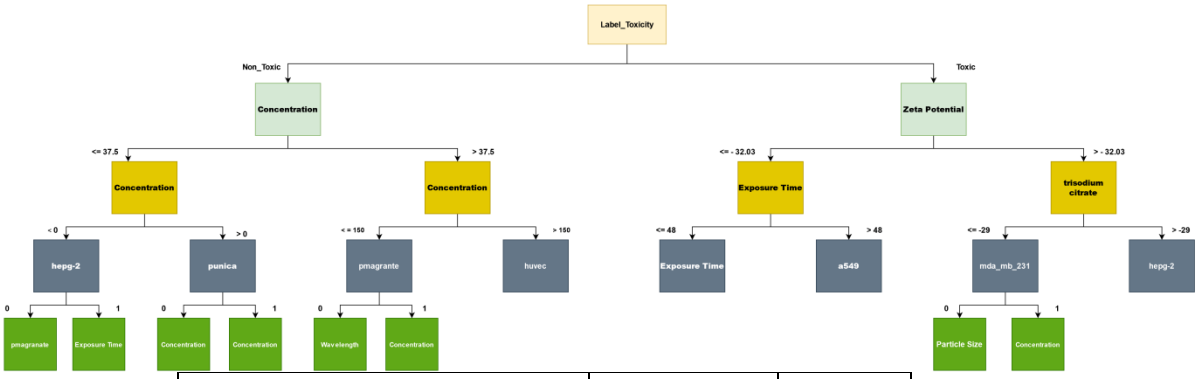
The information is gathered and shaped into an m×n matrix, where m stands for 1135 datasets and n for the 11 characteristics.The work was divided into two sets based on cell lines, such as normal and cancer. Every time identical input features were applied to build both datasets using DT and RF models. The input features were a reducing agent, carcinoma and normal cell lines, exposure period, particle size, hydrodynamic diameter, zeta potential, wavelength, concentration, and cell viability.The input features must have the best possible correlation. As a result, while completing the regression and classification analysis through DT, RF, and clustering, the projected outcome must comply with the conditions involved. Before making a prediction, supervised learning techniques follow a set process. Datasets are divided into training and test datasets when using supervised machine learning. To match the result, algorithms are first trained on the labelled dataset (in terms of cytotoxicity). Second, a subset of the training dataset called the test dataset is used to validate the model. For the best possible anticipated outcome, the input features are chosen. The Orange open-source toolkit's Decision Tree (DT) and Random Forest (RF) were zutilized [56]. Certain restrictions applied to the normal cell line, including the following: (i) cell viability of less than or equal to 50% suggests that nanoparticles are harmful to cells(ii) Cell viability of more than 50% indicates that nanoparticles are not toxic to cells.

The optimal cart segmentation procedure is chosen after repeatedly sampling the original data. At each sampling, a collection of features from each node pool is randomly selected. Finally, a forest is built by aggregating the random features (classifiers) and letting each tree decide which classification is more likely. RF models are often more accurate and resilient than DT classifiers regarding noise and outliers.The conditions imposed on the dataset during the CART algorithm are if the cell viability is greater than or equal to 50%, the nanoparticles are toxic to the cell lines irrespective of whether it is normal or carcinoma, and

if the cell viability is greater than 50%, the nanoparticles are non-toxic to the cell lines irrespective of whether they belong to normal or carcinoma. The root mean square error (RMSE) and R^2 metrics were used to evaluate the performance of the model in regression analysis. RMSE measures the standard deviation of the residuals. R^2 quantifies how well a regression model fits a dataset and measures how well the model reproduces the observed results based on the proportion of the overall variation in the outcomes the model is accountable for explaining.When the regression analysis was performed on the decision tree and random forest model, the R^2 (0.9761) of the DT model was higher than R^2 (0.8776) of the RF model, and the RMSE (3.9322) value of DT was lower than the RMSE (8.8889) value of RF. The high value of R^2 and low value of RMSE indicates that the prediction is accurate, suggesting that the decision tree performed better than the random forest in predicting the toxicity parameter.

Figure 2. Prediction of cell viability using a DT using toxicity as an important parameter for normal cell lines

Table 1.Summary of the selected input features and machine learning results



Model	RMSE	R^2
Decision Tree (DT)	3.9	0.98
Random Forest (RF)	8.9	0.88

The Gini impurity in the decision tree algorithm represents a function that determines how successfully the decision tree was split, and its value ranges between 0 and 0.5. The significance of the chosen input features was ranked using the permutation-based measure called mean decrease Gini (MDG), where the inclusive parameter toxicity was selected since the conditions on which the DT and RF are to be run depends upon the cell viability, which is derived from the cell type. Cell viability was found to be a vital predictor for studying the toxic nature of nanoparticles. Hence, the input feature silver seems to be another potential predictor and the most important determinant of cytotoxicity.

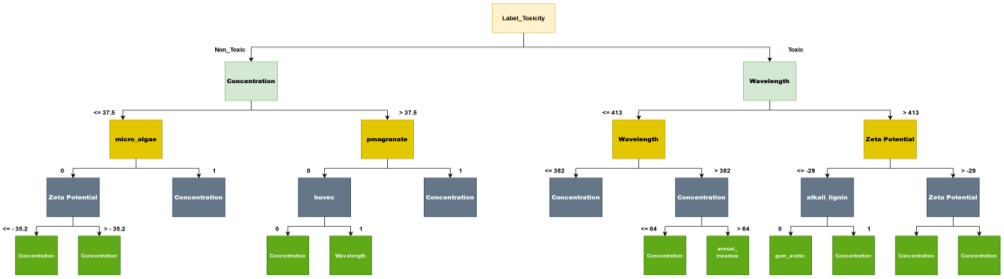


Figure 3. DT confirms that while calculating cytotoxicity and cellviability for carcinoma cell lines, zeta potential, exposure time, reducing agent, and concentration are important factors

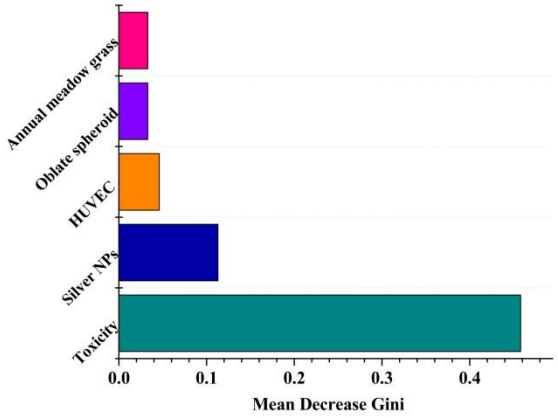


Figure 4. Ranking of the variables using mean decrease Gini for normal cell lines

Table 2. Ranking of the variables using mean decrease Gini for normal cell lines

Parameters	**Gini coefficient**
Toxicity	0.458
Silver NPs	0.113
HUVEC	0.046
Oblate spheroid	0.033
Annual meadow grass	0.032

Table II shows that toxicity is important in combination with Ag-NPs, exposure time, oblate spheroid shape and extract of annual meadow grass (since 162 data points have annual meadow grass as the reducing agent). Thus by these parameters, it can be predicted whether particular materials are toxic or non-toxic to normal cell lines.

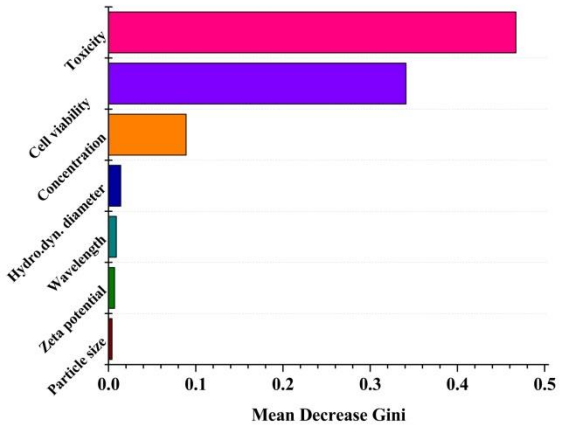


Figure 5. Ranking of the variables using mean decrease Gini for carcinoma cell lines

Table 3. Ranking of the variables using mean decrease Gini for carcinoma cell lines

Parameters	**Gini coefficient**
Toxicity	0.467
Cell viability	0.341
Concentration	0.089
Hydrodynamic diameter	0.014
Wavelength	0.009
Particle size	0.007
Zeta potential	0.004

Table III shows that concentration is an important factor than particle size and zeta potential that affect cell viability in the case of carcinoma cell lines.

4. CONCLUSION

From thecurrentmeta-analysis study, it has been revealedthat the cytotoxicity of Ag-NPs depends upon the physical input features like reducing agents, particle size, zeta potential, cell type (cancer/normal cell lines), hydrodynamic diameter, wavelength, morphology, exposure time and exposure dosage through two well-known supervised machine learning algorithms for regression analysis such as decision tree (DT) and random forest (RF) and the obtained test scores were compared with the DT showing an optimum accuracy in

comparison to RF. The obtained high value of R^2 and low value of RMSE indicated that the prediction is accurate, suggesting that DT has performed better than RF in predicting the toxicity parameter.The forecastis more precise and accurate, and best fits the dataset.The plant-mediated and non-plant mediated cytotoxicity on Ag-NPs with the inclusion of other major variables like cell assays, biological extracts and reaction parameters utilized in synthesis and the performance of models can be further advanced in terms of different machine learning algorithms like unsupervised k-means clustering that can be built to classify so that it will provide a basis for future work.

5. REFERENCES

[1] Feldheim, D. and Keating, C. Self-assembly of single electron transistors and related devices. Chemical Society Reviews, 27(1), pp.1-12, 1998.

[2] Faraday, M., LIX. Experimental relations of gold (and other metals) to light.—The bakerian lecture. The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, 14(96), pp.512-539, 1857.

[3] Gupta, A. and Silver, S. Molecular genetics: silver as a biocide: will resistance become a problem?. Nature biotechnology, 16(10), pp.888-888, 1998.

[4] Choudhury, A. Polyaniline/silver nanocomposites: Dielectric properties and ethanol vapour sensitivity. Sensors and Actuators B: Chemical, 138(1), pp.318-325, 2009.

[5] Klasen, H.J. A historical review of the use of silver in the treatment of burns. II. Renewed interest for silver. Burns, 26(2), pp.131-138, 2000.

[6] Saint, S., Elmore, J.G., Sullivan, S.D., Emerson, S.S. and Koepsell, T.D. The efficacy of silver alloy-coated urinary catheters in preventing urinary tract infection: a meta-analysis. The American journal of medicine, 105(3), pp.236-241, 1998.

[7] Totaro, P. and Rambaldini, M. Efficacy of antimicrobial activity of slow release silver nanoparticles dressing in post-cardiac surgery mediastinitis. Interactive CardioVascular and Thoracic Surgery, 8(1), pp.153-154, 2009.

[8] Wang, L., Hu, C. and Shao, L. The antimicrobial activity of nanoparticles: present situation and prospects for the future. International journal of nanomedicine, 12, p.1227, 2017.

[9] Hebeish, A., El-Rafie, M.H., El-Sheikh, M.A., Seleem, A.A. and El-Naggar, M.E. Antimicrobial wound dressing and anti-inflammatory efficacy of silver nanoparticles. International journal of biological macromolecules, 65, pp.509-515, 2014.

[10] Panáček, A., Kolář, M., Večeřová, R., Prucek, R., Soukupová, J., Kryštof, V., Hamal, P., Zbořil, R. and Kvítek, L. Antifungal activity of silver nanoparticles against Candida spp. Biomaterials, 30(31), pp.6333-6340, 2009.

[11] Mason, C., Vivekanandhan, S., Misra, M. and Mohanty, A.K., Switchgrass (Panicum virgatum) extract mediated green synthesis of silver nanoparticles. World Journal of Nano Science and Engineering, 2(2), p.47, 2012.

[12] Mafuné, F., Kohno, J.Y., Takeda, Y., Kondow, T. and Sawabe, H. Formation

and size control of silver nanoparticles by laser ablation in aqueous solution. The Journal of Physical Chemistry B, 104(39), pp.9111-9117, 2000.

[13] Zhu, J., Liu, S., Palchik, O., Koltypin, Y. and Gedanken, A. Shape-controlled synthesis of silver nanoparticles by pulse sonoelectrochemical methods. Langmuir, 16(16), pp.6396-6399, 2000.

[14] Desai, A.S., Singh, A., Edis, Z., Haj Bloukh, S., Shah, P., Pandey, B., Agrawal, N. and Bhagat, N. An In Vitro and In Vivo Study of the Efficacy and Toxicity of Plant-Extract-Derived Silver Nanoparticles. Journal of Functional Biomaterials, 13(2), p.54, 2022.

Biography

Corresponding Author: Dr. Neeru Bhagat

Dr. Neeru Bhagat is an Associate Professor in the Department of Applied Sciences. She has completed her BSc. and MSc. from Devi Ahilya Vishva Vidyalaya (DAVV) Indore and Ph.D. in Physics from Inter University Consortium, UGC-DAE under DAVV Indore. She is currently working on synthesis and characterization of metal nanoparticles and exploring their biological applicability. She is also working in Heusler alloys as half-metallic ferromagnets for shape memory alloys. She has published more than 25 papers in reputed journals.