On Cytotoxicity Classification of Nanoparticles Via Regularized Regression Models

Aries P. Valeriano, Farlley G. Bondaug, Irelie P. Ebardo, Phoebe A. Sombilon, Gabriel C. Engcong, John Riz V. Bagnol, Mary Joy R. Latayada, Jay Michael R. Macalalag, Brian Dominic J. Paradero, Maricris L. Mayes, Mannix P. Balanay, Arnold C. Alguno and Rey Y. Capangpangan

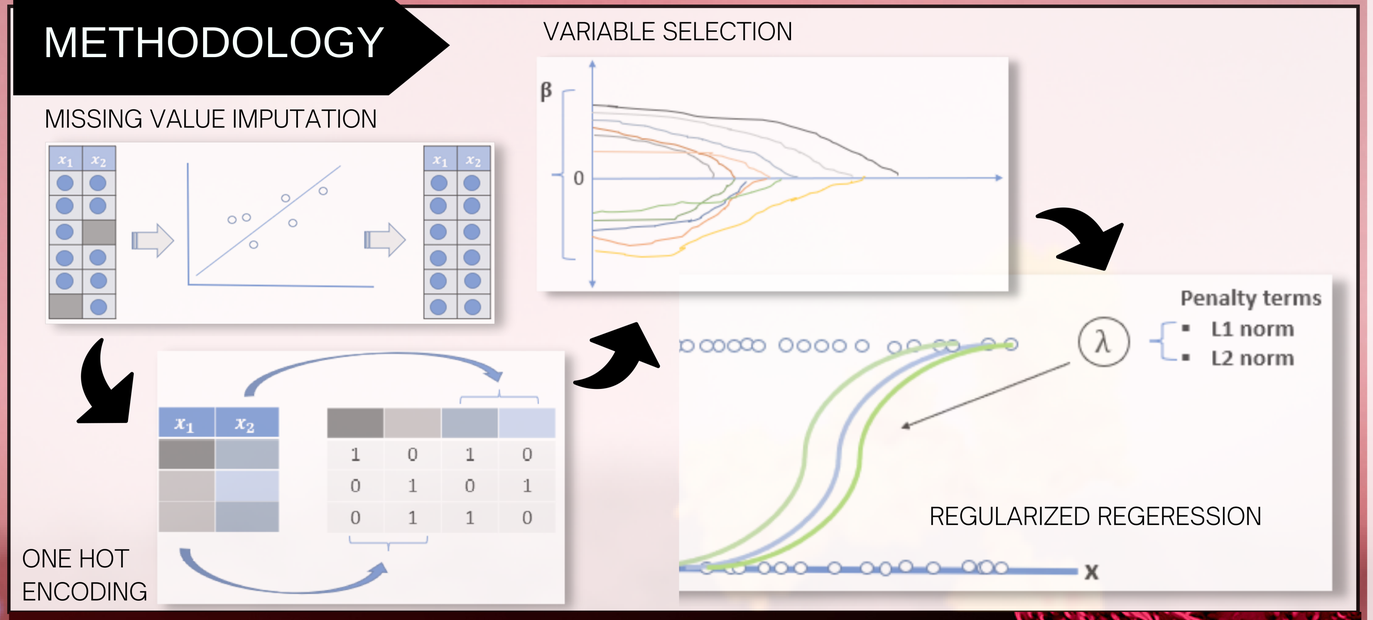
ABSTRACT

The application of engineered nanoparticles (NPs) in various industries, such as pharmaceutical, manufacturing, and food, among others, has proven to be effective through the years. However, NPs are known to produce toxicological effects upon modification of their physicochemical and structural properties. Thus, understanding NPs toxicity behavior is of great importance. Currently, existing studies evaluating the toxicity of different NPs mostly involve employing in vivo (within living organisms) and in vitro (on test tubes or petri dishes) techniques, notwithstanding, these are effective but inefficient, bringing us to the development of the in silico technique (use of a computer or mathematical models). As such, this paper adapts in silico to predict cytotoxicity of different engineered NPs. The dataset used is the one compiled by the group of Gul (2021), consisting of 4111 instances, 25 predictors that composed of physicochemical, test-related, and cell-related properties, with cytotoxicity as its target variable labeled as "high" if (% cell viability ⩽ 50) and "low" if (% cell viability > 50). Upon data exploration and pre-processing, regularized regressions such as ridge, LASSO, and elastic net classification models attain 91.14%, 91.42%, and 91.34% F1 scores at optimal penalty parameter λ's: 0.012, 0.003, and 0.003, respectively in the training set. Moreover, after external validation, the 3 models achieved F1 scores which are as follows: 90.28%, 90.53%, and 90.12%, respectively. These indicates that the built models are able to generalized well to unseen data and are highly accurate in classifying cytotoxicity levels. Additionally, cell type, particularly BEC is the most important variable generally identified by the 3 models.

INTRODUCTION

- Nanotechnology involves understanding and manipulation of matter at the nanoscale level (with length scale 1-100 nanometers) [2].
- NPs are known to produce toxicological effects biologically and environmentally.
- *In vivo* and *in vitro* are commonly used approaches to evaluate the toxicity of NPs, but are unethical and costly.
- *In silico* approach is an alternative to *in vivo* and *in vitro approaches* that is quite popular nowadays as it showed great performance so far.
- This study adapts *in silico,* employing regularized regression such as Ridge, LASSO, and Elastic net to predict cytotoxicity of several engineered NPs.

METHODOLOGY



RESULTS

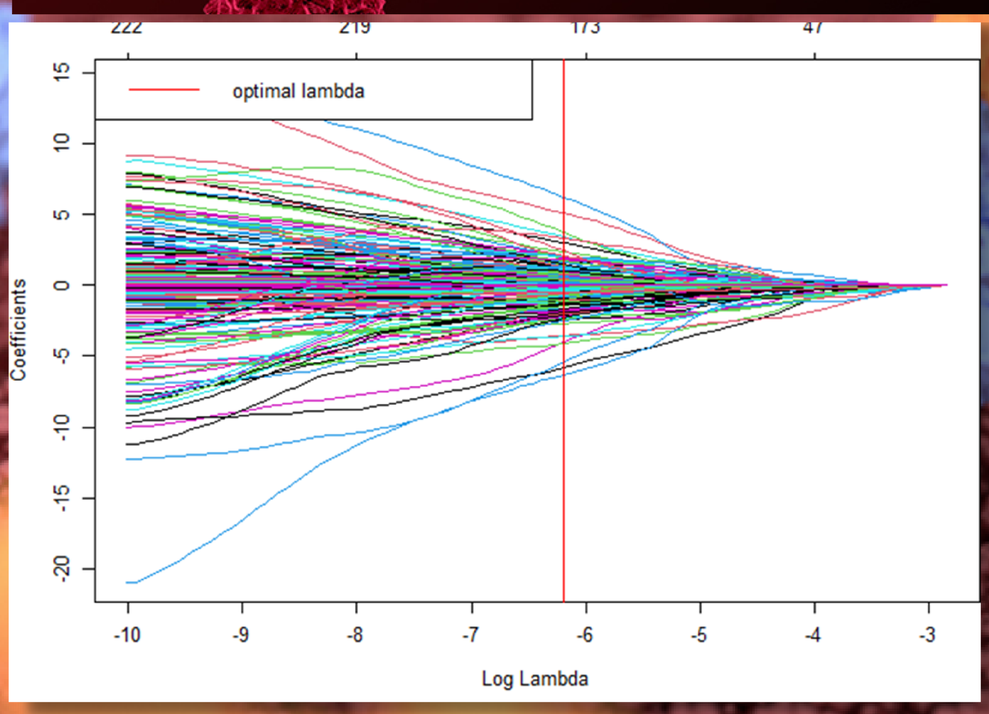


Figure 2. Selection of Influential Variables to Cytotoxicity

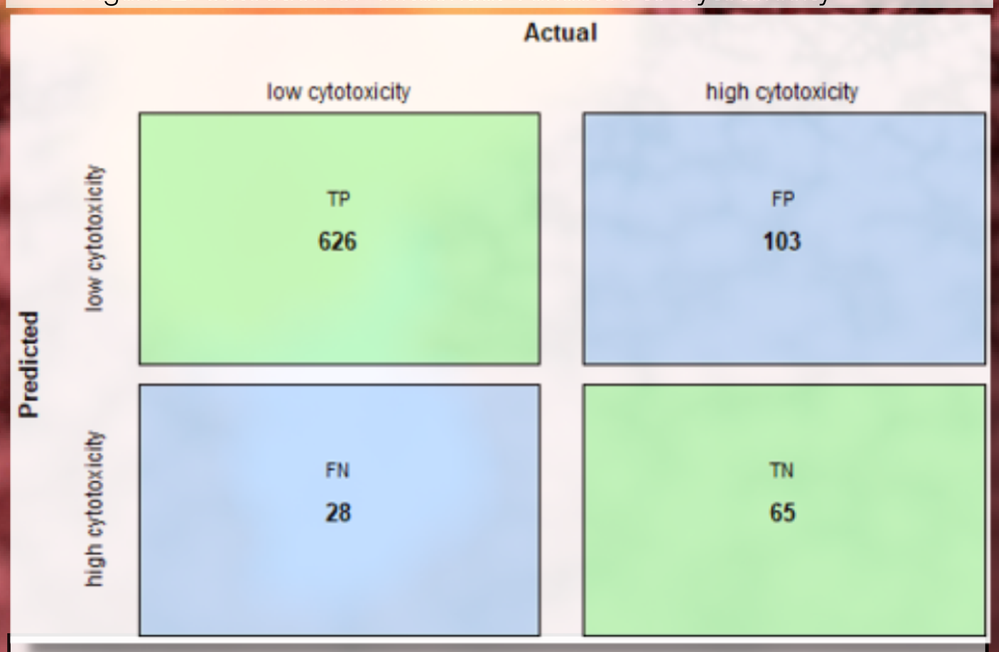


Figure 3. Confusion Matrix using LASSO regression model

RESULTS

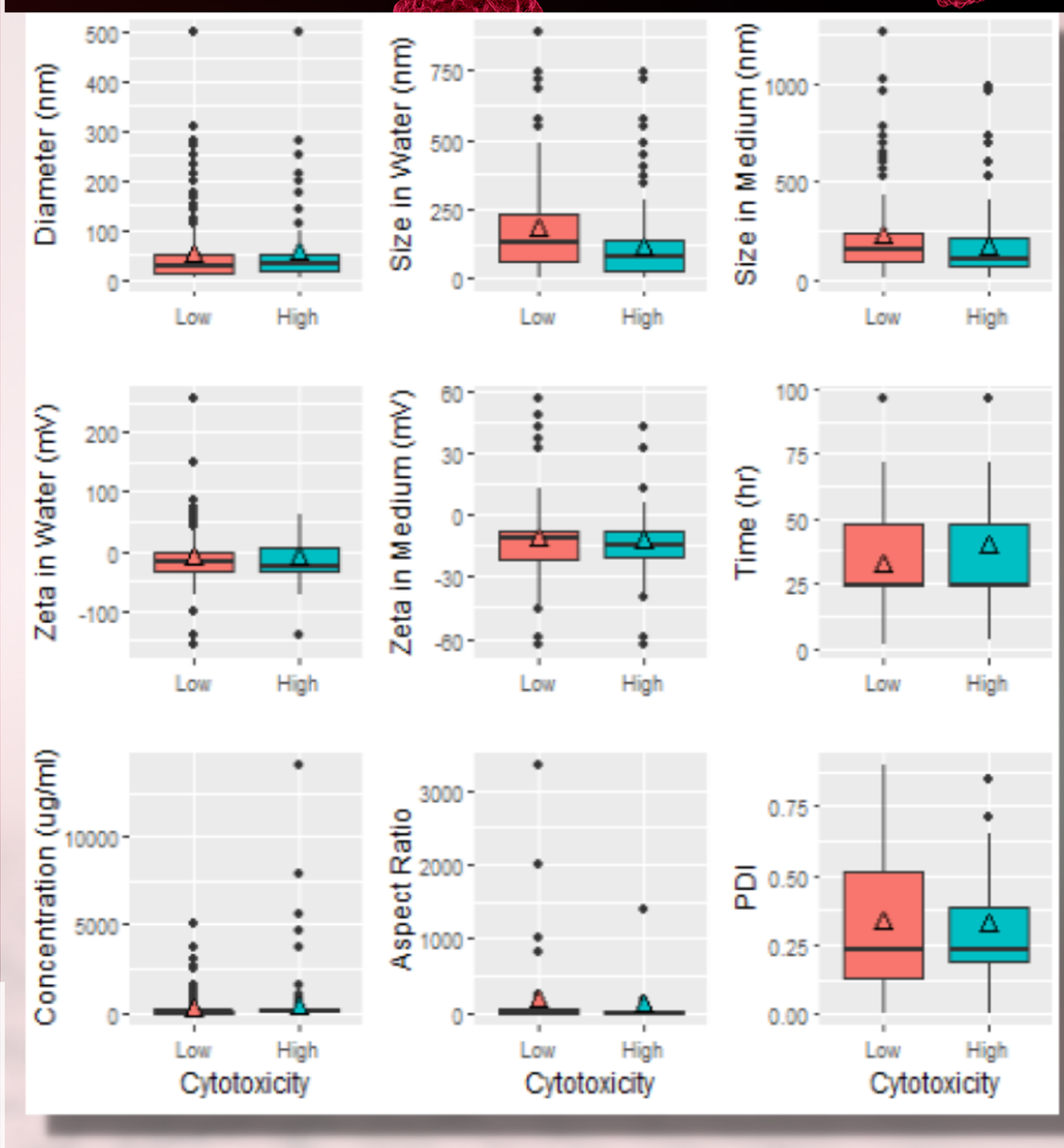


Figure 1. NPs Cytotoxicity Class Distribution by Quantitative Property: high (% cell viability <= 50) or low (% cell viability > 50)

Table 1. Model Evaluation

Models	Optimal λ	F1 score (Training)	F1 score (Testing)
Ridge	0.01231	91.14%	90.28%
LASSO	0.00271	91.42%	90.53%
Elastic Net	0.00311	91.34%	90.12%

CONCLUSION AND RECOMMENDATION

- The 3 regularized regression models, namely, Ridge, LASSO, and Elastic Net all achieved an F1 score greater than both in internal (training) and external (testing) validation, implying that these models are highly accurate in classifying dichotomous levels of NPs cytotoxicity and are able to generalized well to unseen data.
- In addition, these 3 predictive models in general determine cell type, particularly BECs or human bronchial epithelial cells as the most significant variable influencing NPs cytotoxicity. This makes sense as one of the factors affecting the toxicity profile of NPs is their interaction with biological molecules. This further gives an insight that cytotoxicity is clearly determined if the route of entry is through inhalation.
- For further study, combine the type of cell lines based on what routes of entry it belongs to, namely, digestion, skin penetration, inhalation, and injection during medical procedures among others. Generally, this will allow us to see if cytotoxicity can be identified clearly through one of these entries.

REFERENCES

1. Gul, G., Yildirim, R., & Ileri-Ercan, N. (2021). Cytotoxicity analysis of nanoparticles by association rule mining. Environmental Science: Nano, 8(4), 937-949. https://doi.org/10.1039/d0en01240h
2. What Is Nanotechnology? | National Nanotechnology Initiative. Retrieved October 20, 2022, from https://www.nano.gov/nanotech-101/what/definition