



## 05. NIXMSM-07: North-Ibero-American Exp., Model., and Simul. Methods Congress, Valencia, Spain-Miami, USA, 2021



### Critical essay on predictive models for anti-sarcoma compounds

Bernabe Ortega-Tenezaca <sup>a, b</sup>

<sup>a</sup> RNASA-IMEDIR, Computer Science Faculty, University of A Coruña, 15071, A Coruña, Spain.

<sup>b</sup> Universidad Estatal Amazónica UEA, Puyo, Pastaza, Ecuador.

#### Abstract.

Today, studies are performed from a dataset spanning multiple preclinical assays and different experimental conditions for sarcomas. PTML is a tool that combines Machine Learning (ML) algorithms and Perturbation Theory (PT) principles. With PTML, ML techniques can be used to predict antisarcoma compounds. At the same time, different PT techniques can be applied. One of the most widely used ML techniques is the neural network which showed high accuracy for both training and model validation. It is important to emphasize that the production of the most optimal model would save resources in the pharmaceutical industries. In a recent paper Cabrera *et al.* reported a new model for prediction of anti-sarcoma compounds. The model is very interesting because it can predict the biological activity vs multiple proteins, etc. The authors also explored multiple molecular descriptors of drugs as well as many assay conditions like protein target, cell line, etc. There are some suggestions we can make to improve future versions of this paper. For instance, the authors could calculate also sequence descriptor of target proteins to predict the results for new mutants. On my opinion, it could be very interesting developing a user-friendly software for use of non-expert medicinal chemists. This software could be a desktop or online server application increasing the use of the model worldwide. Another interesting step could be the fusion of the present pre-clinical data with clinical data including variables of patients or population groups. In all case, the paper is very interesting an opens new gates to the authors for future works including new features to the design of antisarcoma compounds.

## References

1. Cabrera-Andrade, A.; López-Cortéz, A.; Munteanu, C.; Pazos, A.; Pérez-Castillo, Y.; Tejera, E.; Arrasate, S.; González-Díaz, H. Perturbation-Theory Machine Learning (PTML) Multilabel Model of the ChEMBL Dataset of Preclinical Assays for Antisarcoma Compounds. *ACS Omega*. 2020, 5, 42, 27211-27220.
2. Gaulton, A.; Hersey, A.; Nowotka, M.; Bento, A. P.; Chambers, J.; Mendez, D.; Mutowo, P.; Atkinson, F.; Bellis, L. J.; Cibrián-Uhalte, E.; Davies, M.; Dedman, N.; Karlsson, A.; Magariños, M. P.; Overington, J. P.; Papadatos, G.; Smit, I.; Leach, A. R. The ChEMBL database in 2017. *Nucleic Acids Res.* 2017, 45, D945–D954.
3. Lo, Y.-C.; Rensi, S. E.; Torng, W.; Altman, R. B. Machine learning in chemoinformatics and drug discovery. *Drug Discovery Today* 2018, 23, 1538–1546.
4. Singh, H.; Kumar, R.; Singh, S.; Chaudhary, K.; Gautam, A.; Raghava, G. P. S. Prediction of anticancer molecules using hybrid model developed on molecules screened against NCI-60 cancer cell lines. *BMC Cancer* 2016, 16, 77.
5. Toropov, A. A.; Toropova, A. P.; Benfenati, E.; Gini, G.; Leszczynska, D.; Leszczynski, J. SMILES-based QSAR approaches for carcinogenicity and anticancer activity: comparison of correlation weights for identical SMILES attributes. *Anti-Cancer Agents Med. Chem.* 2011, 11, 974–982.
6. González-Díaz, H.; Bonet, I.; Terañ, C.; De Clercq, E.; Bello, R.; García, M. M.; Santana, L.; Uriarte, E. ANN-QSAR model for selection of anticancer leads from structurally heterogeneous series of compounds. *Eur. J. Med. Chem.* 2007, 42, 580–585.
7. González-Díaz, H.; Gia, O.; Uriarte, E.; Hernández, I.; Ramos, R.; Chaviano, M.; Seijo, S.; Castillo, J. A.; Morales, L.; Santana, L.; Akpaloo, D.; Molina, E.; Cruz, M.; Torres, L. A.; Cabrera, M. A. Markovian chemicals "in silico" design (MARCH-INSIDE), a promising approach for computer-aided molecular design I: discovery of anticancer compounds. *J. Mol. Model.* 2003, 9, 395–407.
8. Jung, M.; Kim, H.; Kim, M. Chemical genomics strategy for the discovery of new anticancer agents. *Curr. Med. Chem.* 2003, 10, 757–762.
9. Shi, L. M.; Fan, Y.; Myers, T. G.; O'Connor, P. M.; Paull, K. D.; Friend, S. H.; Weinstein, J. N. Mining the NCI anticancer drug discovery databases: genetic function approximation for the QSAR study of anticancer ellipticine analogues. *J. Chem. Inf. Comput. Sci.* 1998, 38, 189–199.
10. Yang, J.; Ren, Z.; Du, X.; Hao, M.; Zhou, W. The role of mesenchymal stem/progenitor cells in sarcoma: update and dispute. *Stem Cell Investig.* 2014, 1, 18.