

## [P14] Probing the Hypothesis of SAR Continuity Restoration by Activity Cliffs Removal in QSAR Modeling

Maykel Cruz-Monteagudo<sup>1,2,3</sup>, Yunier Perera-Sardiña<sup>3</sup>, Yunierkis Pérez-Castillo<sup>3</sup>, Fernanda Borges<sup>1</sup>, José L. Medina-Franco<sup>4</sup>, Orazio Nicolotti<sup>5</sup>, M. Natália D. S. Cordeiro<sup>2</sup>

<sup>1</sup>CIQ, Department of Chemistry and Biochemistry, Faculty of Sciences, University of Porto, 4169-007 Porto, Portugal.

<sup>2</sup>REQUIMTE, Department of Chemistry and Biochemistry, Faculty of Sciences, University of Porto, 4169-007 Porto, Portugal.

<sup>3</sup>CEQA, Faculty of Chemistry and Pharmacy & Molecular Simulation and Drug Design Group, CBQ, Central University of Las Villas, Santa Clara 54830, Cuba.

<sup>4</sup>Mayo Clinic, 13400 East Shea Boulevard, Scottsdale, AZ 85259, USA.

<sup>5</sup>Dipartimento di Farmacia-Scienze del Farmaco, Università degli Studi di Bari 'Aldo Moro', Via Orabona 4, 70125 Bari, Italy.

In this work we report the first attempt to study the effect of activity cliffs over the generalization ability of machine learning models, using as study case a previously reported diverse dataset focused on drug induced liver injury (DILI). Here, the hypothesis of SAR continuity restoration by activity cliffs removal is tested as a potential solution to overcome such limitation. Previously, a parallelism was established between activity cliffs and instances that should be misclassified (ISMs), a related concept from the field of machine learning [1, 2]. Based on this concept we comparatively studied the classification performance of multiple machine learning classifiers as well as the consensus classifier derived from predictive classifiers obtained from training sets including or excluding ISMs. The influence of the removal of ISMs from the training set over the virtual screening performance was also studied for the respective consensus classifiers algorithms. In comparison with classifiers obtained from the unaltered DILI training set (RAW-Classifiers), those classifiers trained after ISM removal (ISM-Classifiers) do not exhibited a statistically significant average difference of accuracy on the external evaluation set (0.75%) but a statistically significant deterioration on specificity (-1.65%). The sensitivity, the weakest feature of RAW-Classifiers was significantly improved by ISM-Classifiers in 3.73%. A similar trend was observed for the consensus classifiers derived from the respective acceptable RAW- and ISM-Classifiers with a more noticeable difference, favoring RAW-Classifiers. However, analyzing the virtual screening performance of each consensus classifier was observed a noticeable difference favoring ISM-Classifiers. Although the improvement on overall enrichment was not marked there was a significant improvement on initial enrichment, especially at early fractions (1% and 5%) [3]. The improvement on virtual screening performance of consensus ISM-classifiers can be explained by the improvement on sensitivity, the weakest feature of RAW-Classifiers. These results suggest a positive effect of ISM removal for practical applications. Even when a significant improvement is not attained on classification accuracy, the virtual screening performance of ISM-Classifiers is positively influenced, significantly improving their early recognition ability.

[1] Cruz-Monteagudo, M.; Medina-Franco, J. L.; Perez-Castillo, Y.; Nicolotti, O.; Cordeiro, M. N.; Borges, F. Drug Discov. Today (2014). In press. DOI : 10.1016/j.drudis.2014.02.003

[2] Smith, M. R.; Martinez, T. The 2011 International Joint Conference on Neural Networks (2011) 2690-2697.

[3] Perez-Castillo, Y.; Cruz-Monteagudo, M.; Lazar, C.; Taminau, J.; Froeyen, M.; Cabrera-Perez, M. A.; Nowe, A. Mol. Divers. (2014) In press. DOI: 10.1007/s11030-014-9513-y.