

PTMLIF MODEL OF ChEMBL PRECLINICAL ASSAYS OF VIT D DERIVATIVES
VS.
SINGLE NUCLEOTIDE POLYMORPHISM (SNP) DATA

Cristina Lanza

Abstract:

The vitamin D receptor is a common target for various drugs, and is of great interest due to the protective function that this vitamin exerts on the body. The presence of single nucleotide polymorphisms (SNPs) in this receptor can affect the binding of the drug, which makes its analysis important. Through chemoinformatic studies, Perturbation Theory Machine Learning Information Fusion (PTMLIF) models that analyze this interaction can be established, for which the drug data set was downloaded from the public databases ChEMBL and NCBI (National Center for Biotechnology Information) and was subsequently performed the fusion of information. The database included 26064 trials with 47 different properties and 376 SNPs. In the present study, the deviations of the reference drug with respect to the perturbation operators were measured. The best model obtained showed values of $Sp = 72.62\%$, $Sn = 89.54\%$ and $Ac = 83.85\%$ for training and $Sp = 74.78\%$, $Sn = 90.88\%$ and $Ac = 85.31\%$ for validation for a given application domain.

References:

- Bediaga, H., Arrasate, S., & González-Díaz, H. (2018). PTML Combinatorial Model of ChEMBL Compounds Assays for Multiple Types of Cancer. *ACS Combinatorial Science*, 20(11), 621–632. <https://doi.org/10.1021/acscombsci.8b00090>
- Gonzalez-Diaz, H. (2018). *PTML: Perturbation-Theory Machine Learning notes*. 5463. <https://doi.org/10.3390/mol2net-04-05463>
- Belhassan, A., Chtita, S., Lakhlifi, T., & Bouachrine, M. (2018). *QSRR Study of Linear Retention Indices for Volatile Compounds using Statistical Methods*. 7(4), 558–575. <https://doi.org/10.7598/cst2018.1501>