*Prediction of the n-octanol/water distribution coefficient (log D) at different pH values by a hybrid Machine Learning/Physical method.*

The distribution coefficient between *n*-octanol/water (log *D*) is a lipophilicity descriptor widely used in medicinal chemistry.  Most of the scientific literature reports it or predicts it at physiological pH (7.4). In this work we integrated the prediction of the partition coefficient of the neutral species (log $P_N$), the partition coefficient of the ionized species (log $P_I$) and the acid dissociation constant (p$K_a$) by Machine Learning algorithms; to finally get an equation that describes the distribution coefficient in function of the pH. The *Random Forest* algorithm was used to predict the log $P_I$ and *XGBoost* for the log $P_N$ and p$K_a$. In our test set (n = 298), containing values in the pH range of 1.0-13.0, we successfully predicted the *log D* with a RMSE of 0.76 *log D* units. Further evaluation, with the SAMPL7 set (n = 22, pH = 7.4), was done and a RMSE of 0.96 *log D* units was achieved. These results show a comparable accuracy with typically used licensed software predictions.