Detección de lenguaje ofensivo con redes neuronales profundas

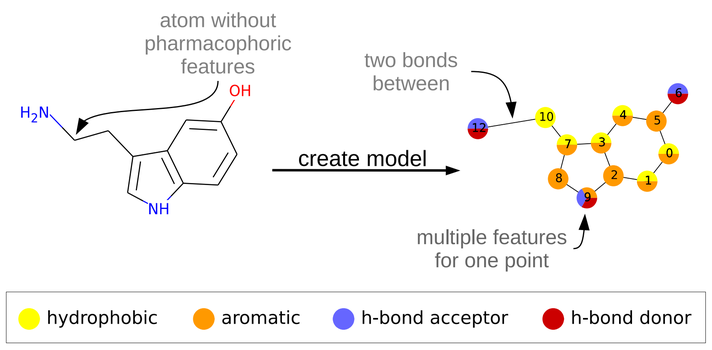
Edson Raul Cepeda Marquez
Facultad de Ingeniería Mecánica y Eléctrica
raulcepedac@hotmail.com

1. Introducción

Uno de los retos actuales de internet es el de mantener las plataformas digitales libres de agresiones y promover un ambiente sano para los usuarios. Con el uso de redes sociales las personas expresan sus opiniones llegando a esparcir mensajes de odio, insultos o discriminación. El objetivo de este proyecto es identificar correctamente texto con lenguaje ofensivo utilizando técnicas de procesamiento de lenguaje natural y redes neuronales profundas.

2. Pharmacophore model

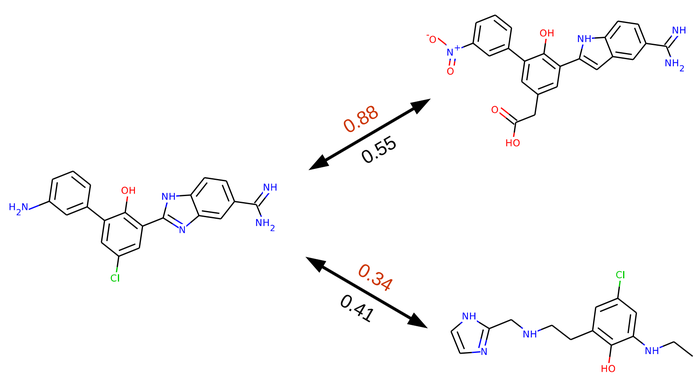
To describe a molecule, DeCAF substitutes its functional groups with pharmacophoric points (hence the "F" in the algorithm's name). Points are organised into an undirected graph. Lengths of the edges in the graph represents the number of bonds between pharmacophoric points.



3. Similarity measure

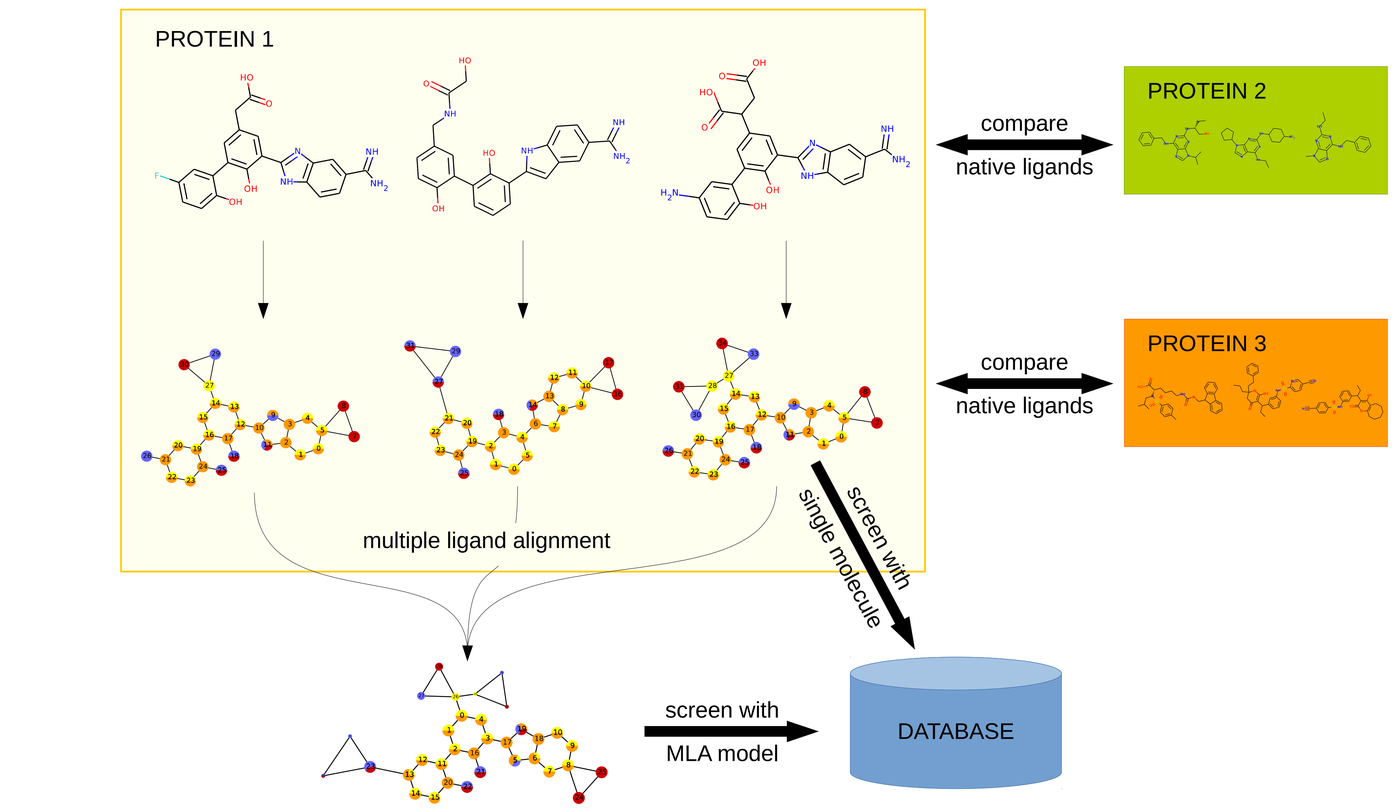
To measure similarity of two molecules or to combine them into one model, DeCAF first finds their **maximum common substructure (MCS)**. To provide fast, but accurate method for solving MCS problem, we combined Generic Match Algorithm (GMA) [**?**] with backtracking algorithm proposed by Yiqun Cao [**?**].

Here we present comparison of molecules with similar and with different structures. DeCAF scores and **Tanimoto coefficient (Tc)** values are shown in red and black, respectively.

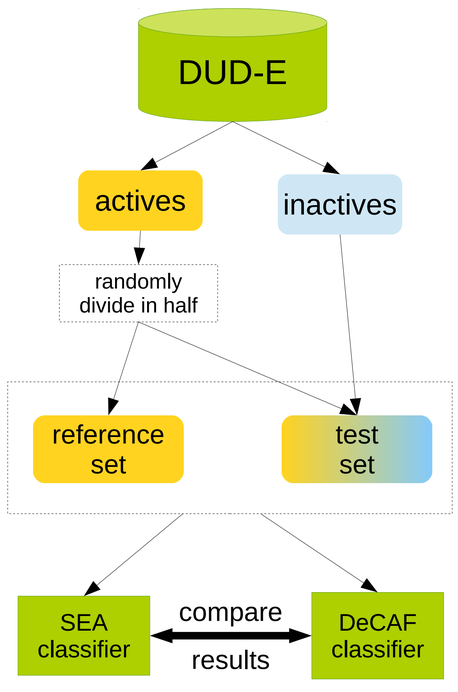


4. Applications

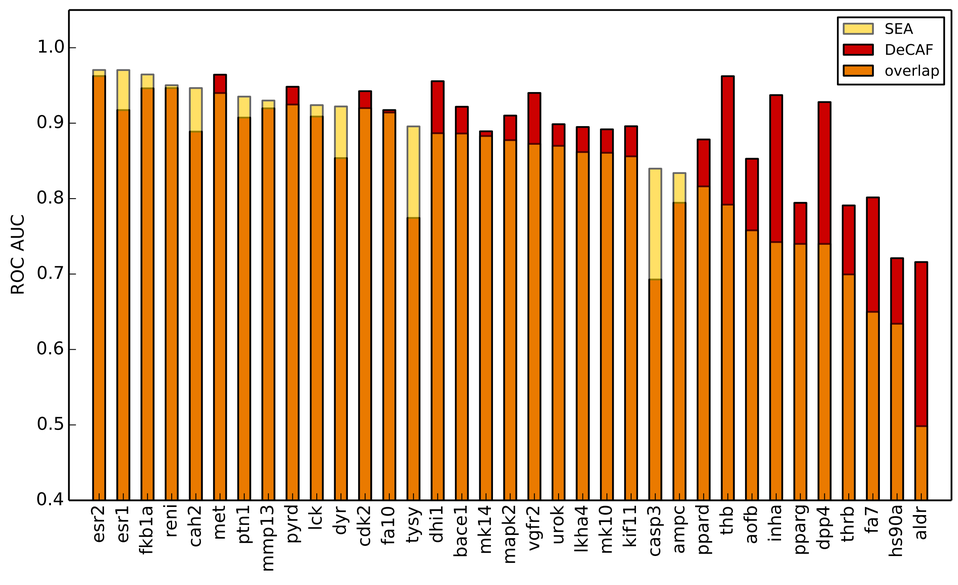
DeCAF is a versatile tool with many possible applications. It allows to compare two molecules or more complex models created from sets of ligands. Our method can be used to align multiple ligands and find crucial pharmacophoric features in a set of active compounds. Pharmacophore models can help in database screening for molecules with desired properties. DeCAF is also suitable for comparing entire sets of ligands, e.g. to analyse properties of proteins in drug repositioning process.



5. DeCAF vs. SEA



We tested DeCAF in 35 diverse targets taken from the DUD-E database, to evaluate its power to classify molecules as active or inactive. We compared DeCAF to the renowned **SEA (Similarity Ensemble Approach)** algorithm [**?**], which uses Tc as a similarity measure. Dataset preparation steps are shown on the left diagram. Comparison results (**ROC AUC** values for each receptor) are shown below.



6. Conclusions

We proved that DeCAF is a significant improvement over the SEA algorithm, a popular method for comparing sets of ligands.

1. DeCAF gives better results for 23 out of 35 receptors.
2. For targets with easily separable active and inactive datasets, SEA and DeCAF give similar results.
3. In cases in which SEA fails to identify active molecules, our method performs substantially better.

7. References