Controlled Conversational Models through Conversation-Dedicated Ontology

Barbara Gendron, under the supervision of Mathieu d'Aquin and Gaël Guibon
Laboratoire Lorrain de Recherche en Informatique et ses Applications (LORIA), Université de Lorraine
barbara.gendron@loria.fr

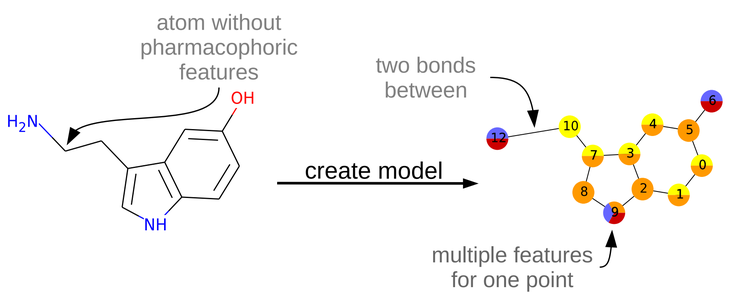
1. Context

Predicting biological activity of small molecules is a key element of computer-aided drug design. Existing methods often fail to identify ligands with similar physicochemical properties but different structures. Many of the current approaches rely on generating 3D conformations, which leads to sampling problems and unacceptably high computational costs for large sets of molecules. Herein we present DeCAF – a novel method for describing ligand properties and a fast and effective tool for comparing multiple molecules, and merging them into a single pharmacophore model. DeCAF is written as an open source Python module (**http://bitbucket.org/marta-sd/decaf**) and can be easily combined with RDKit to facilitate ligand-based drug design.

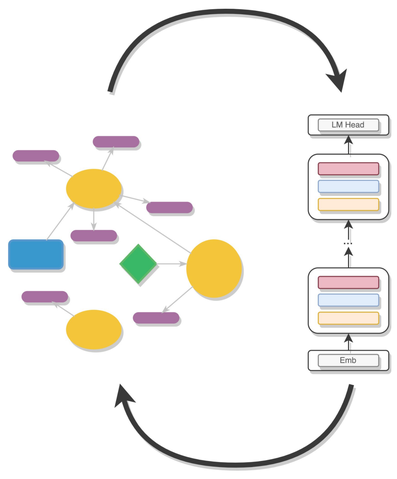
2. Methodology

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. 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. 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. 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. 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.
Some interesting papers:

- One
- Two
- Three



3. Objectives

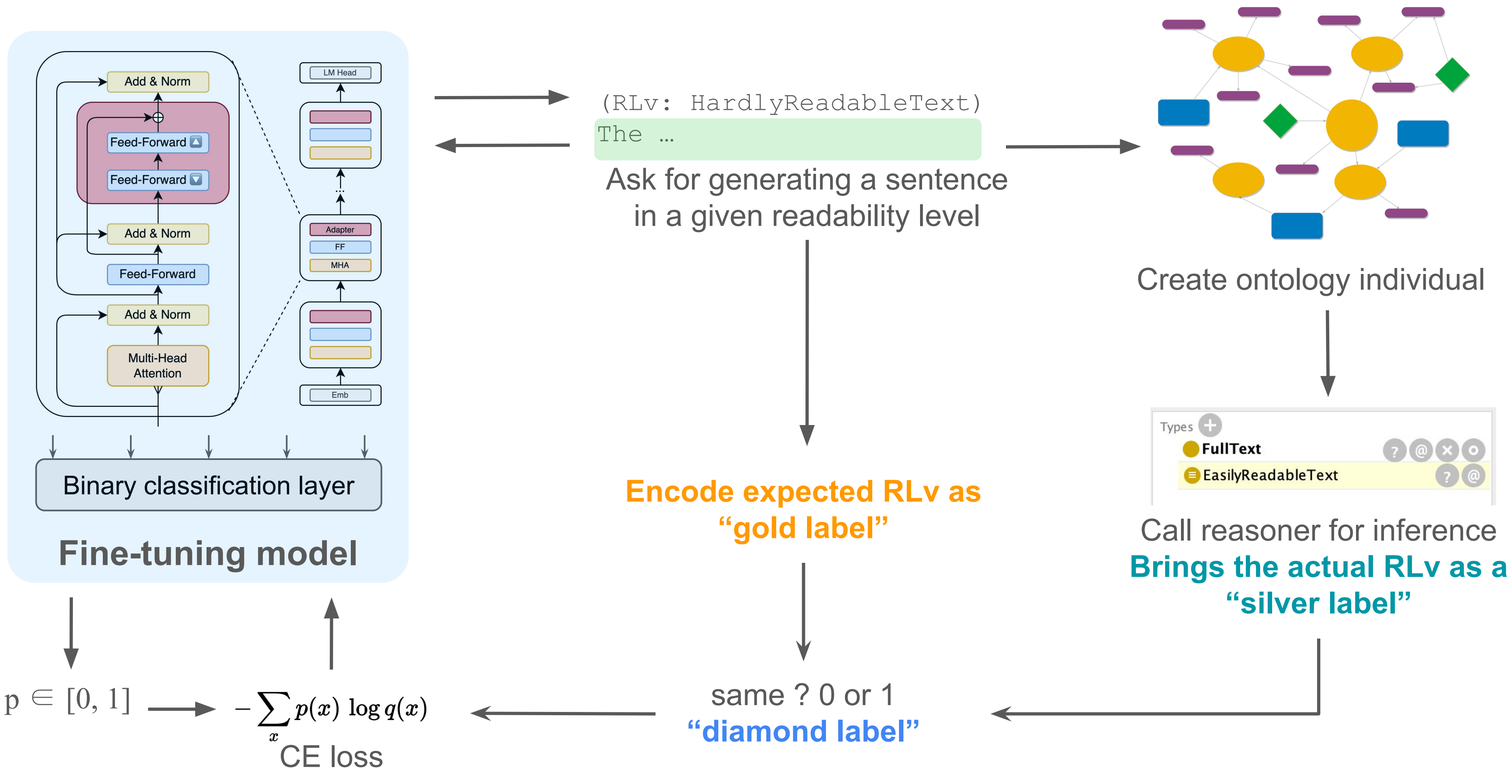


We tested DeCAF in 35 diverse targets taken from the DUD-E database, to evaluate its power to classify molecules as active or inactive. We compared DeCAF to the renowned **SEA (Similarity Ensemble Approach)** algorithm [1], which uses Tc as a similarity measure. Dataset preparation steps are shown on the left diagram. Comparison results (**ROC AUC** values for each receptor) are shown below.

- One
- Two
- Three

4. OntoGPT: LLM Fine-Tuning Based on Ontology Validation

This tool is currently under development...
Blablabla



5. Related Work

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.

6. References

[1] Michael J Keiser, Bryan L Roth, Blaine N Armbruster, Paul Ernsberger, John J Irwin, and Brian K Shoichet. Relating protein pharmacology by ligand chemistry. *Nat. Biotechnol.*, 25(2):197–206, 2007.