Cheminformatics Tutorial Documentation
Release 0.1

Noel M O'Boyle

Nov 01, 2017

Contents

Part 1 - Introduction to Cinfony

17. How do you think does the software convert a name like "aspirin" to a chemical structure? What happens if you try to convert your own name to a chemical structure?

17. What is the structure of piperazine? What is its Fluka catalogue number?

17. Generate a SMILES string for piperidine. Can you think of another SMILES string that would also represent the same molecule? Test whether you are right by reading in that SMILES string and depicting the associated molecule.

17. What molecule has the same SMILES string as piperidine but with all of the letters lowercase?

17. The Wikipedia page for piperazine gives its CAS number, a SMILES string and its InChI. Which of these identifiers is most suitable for searching the web via Google?

CHAPTER 2

Part 2 - Descriptors

The first piece of code in the section on the Rules of Five has an error. Instead of using that, copy the following function (by selecting and then using "CTRL+C") and paste into the Python console (by clicking in the Python console and using "CTRL+V"):

```python
def lipinski(mol):
    descs = [('WeightDescriptor', 'MW'), ('HBondDonorCountDescriptor', 'nHBDon'),(
→'HBondAcceptorCountDescriptor', 'nHBAcc'), ('ALOGPDescriptor', 'ALogp2')]
    targets = [500, 5, 10, 5.0]
    descvalues = mol.calcdesc([x[0] for x in descs])
    failures = 0
    for desc, target in zip(descs, targets):
        descvalue = descvalues["%s_%s" % (desc[0], desc[1])]
        if descvalue > target:
            failures += 1
    return failures
```

17. Look at the structures of the three molecules in the text that fail at least one of the Rule of Fives. How do they compare to typical drug-like molecules?

17. Choose three other common drugs and find out whether they fail the Rule of Fives.

Part 3 - Fingerprints

---

17. Which of the bits of the propanol fingerprint correspond to which of the paths listed? (Assign as many as possible.)

17. In the lecture, we discussed the idea of extending laws governing particular controlled subtances to other molecules with "substantially similar" structure. What is the Tanimoto similarity of the binary fingerprints of cathinone and mephedrone? (Hint: an InChI for mephedrone is available on its Wikipedia page.) Do the same analysis for cathinone and naphyrone (whose SMILES string is `C2CCCN2C(CCC)C(=O)c3cc1ccccc1cc3`).