

使用 MolAICal 和 OnionNet 模型进行 药物亲和力的预测

作者: MolAICal (update 2021-10-16)

更多教程(含英文教程)请见如下:

MolAICal 官方主页: <https://molaical.github.io>

MolAICal 官方主页中国镜像: <https://molaical.gitee.io>

MolAICal 文章介绍: <https://arxiv.org/abs/2006.09747> 和 <https://doi.org/10.1093/bib/bbaa161>

MolAICal 中文博客: <https://molaical.github.io/entutorial.html>

MolAICal Twitter: <https://twitter.com/MolAICal>

MolAICal QQ 学术讨论群: 1151656349

1. 简介

深度学习模型可以用来计算配体和受体之间的亲和力,将深度学习模型 OnionNet 嵌入到 MolAICal 用以计算配体的亲和力。这个案例可以帮助使用者快速了解深度学习是如何计算配体和蛋白的亲和力,有关 MolAICal 的更多详细信息,请参考文章(<https://doi.org/10.1093/bib/bbaa161>)。

2. 软件和所需数据

2.1 软件需求

1) MolAICal: <https://molaical.github.io>

国内镜像 MolAICal: <https://molaical.gitee.io>

2.2 案例文件

2) 所有在教程中的重要文件可以从该网址下载:

<https://gitee.com/molaical/tutorials/tree/master/014-bindingaffinityOnionnet>

3. 步骤

3.1 安装 OnionNet 模型

现阶段只支持 linux 系统

1) 点击并打开网页: [DownloadModel](#)

在已打开的网页中根据路径 `science/AImodels/BindingAffinity/onionnet/linux`, 下载文件名是“onionnet.tar.gz”。

2) 把 `onionnet.tar.gz` 移动到 `MolAICal-xxx/mtools` 目录下, “MolAICal-xxx”是你解压并安装 MolAICal 的目录, “mtools”是指定目录。

3) 解压目录

```
#> tar -xzf onionnet.tar.gz
```

进入到 onionnet 目录下

```
#> cd onionnet
```

4) 安装 OnionNet 模型

```
#> chmod +x install.sh
```

```
#> ./install.sh
```

现在已成功安装 OnionNet 模型

3.2 计算单个复合物的亲和力

3.2.1 将一个蛋白和配体拼接在一起

打开 `014-bindingaffinityOnionNet` 文件夹, 使用以下命令准备复合物

```
#> molaical.exe -model mergeon -r 1a30_protein.pdb -l 1a30_ligand.mol2.pdb -c 1a30_complex.pdb
```

注意: 如果使用者没有 `pdb` 格式的文件, 可以使用 MolAICal 转换分子格式。例如, 如果用户使用的是 `mol2` 格式的文件, 可以使用下面的命令来进行格式转换 (注释: 分子应该由正确的分子后缀, 如 `.mol2`, `.pdb` 等, 否则无法被 MolAICal 自动识别)。

```
#> molaical.exe -tool format -i ligand.mol2 -o ligand.pdb
```

3.2.2 计算 pK_x (pK_a or pK_i) 用于亲和力预测

把合并好的复合物文件放到一个文件夹中, 在这个教程中, 将包含复合物文件名的文件命名为“`inputlist.dat`”, 使用下列命令对 pK_x 进行计算:

```
#> molaical.exe -model onionnet -i inputlist.dat -o results.csv
```

运行结束后, 会生成一个包含 pK_x 值的 `results.csv` 的文件, 如果研究人员想要将 pK_x 值转化为结合自由能, 可以参照该网页 (<https://molaical.github.io/tutorial.html>), MolAICal 提供自由能转化模块。

3.3 计算多个复合物的亲和力

3.3.1 将多个配体与一个蛋白分别对接，形成复合物

有些情况下，研究人员需要计算很多配体的亲和力，基于这种需求，这需要将很多配体与蛋白结合，形成很多复合物。

打开 014-bindingaffinityOnionnet/list 文件夹

将所有配体的名字写入到“list.txt”文件中，在这部分中，蛋白的 pdb 文件是 GCGRNoLigand.pdb。然后使用下列的命令进行蛋白和配体的合并。

```
#> molaical.exe -model mergeon -r GCGRNoLigand.pdb -f list.txt
```

运行完毕后将会生成一个“complex_<your ligand name>.pdb”文件。

3.3.2 计算 pKx (pKa or pKi) 用于亲和力预测

使用如下命令将所有合并的复合物文件名称写到“complexList.txt”:

```
#> ls -A complex_*.pdb > complexList.txt
```

然后使用下列命令进行亲和力计算。

```
#> molaical.exe -model onionnet -i complexList.txt -o results.csv
```

这将生成一个包含所有 pKx 值的 csv 文件(results.csv)，如果研究人员想要将 pKx 值转化为结合自由能，可以参照该网页 (<https://molaical.github.io/tutorial.html>)，MolAICal 提供自由能转化模块。

参考文献:

1. Zheng L, Fan J, Mu Y. OnionNet: a Multiple-Layer Intermolecular-Contact-Based Convolutional Neural Network for Protein-Ligand Binding Affinity Prediction. ACS Omega. 2019;4(14):15956-65. Epub 2019/10/09. doi: 10.1021/acsomega.9b01997. PubMed PMID: 31592466; PubMed Central PMCID: PMC6776976.