



ADMETSCORE

ADMETSCORE	1
Introdução	1
Arquivo .sdf	2
AdmetLab 3.0	3
Instalando o Admetscore	5
Utilização Como Ferramenta	6
Utilização Como Biblioteca	7
Criando Distribuição Binária	7

Introdução

O Admetscore é um software desenvolvido em Python que simplifica e automatiza a comparação de análises ADMET entre várias moléculas. Ele pode ser utilizado tanto como uma ferramenta de linha de comando quanto como uma biblioteca Python.

O programa recebe as moléculas em arquivos .sdf (Structure Data Format), um formato padrão amplamente utilizado para representar estruturas químicas e propriedades associadas. As análises ADMET dessas moléculas são realizadas no site do ADMETlab 3.0, onde os parâmetros são classificados como bons, médios ou ruins com base na avaliação empírica fornecida pelo site. A partir dessa classificação, o Admetscore atribui uma pontuação de 0 a 10 para cada propriedade e grupo ADMET, com 0 representando um desempenho ruim e 10 o melhor desempenho possível. As moléculas são, então, ranqueadas conforme sua pontuação geral, facilitando a identificação das melhores.

O Admetscore gera como saída uma planilha contendo as IDs das moléculas, as SMILES (Simplified Molecular Input Line Entry System), a pontuação total e por grupo ADMET, além de um detalhamento completo da análise ADMET obtida do

ADMETlab 3.0. Cada célula da planilha é codificada por cores—verde, amarelo e vermelho—indicando parâmetros bons, médios e ruins, respectivamente, tornando a interpretação dos resultados mais intuitiva.

Arquivo .sdf

O Structure Data Format (SDF) é um formato amplamente utilizado para armazenar e compartilhar informações sobre estruturas moleculares. Originalmente desenvolvido pela Molecular Design Limited (MDL), o SDF se tornou um padrão para representar moléculas de forma legível tanto por humanos quanto por máquinas. Este formato é especialmente útil para armazenar coordenadas atômicas, conectividade, e outras propriedades químicas de várias moléculas em um único arquivo. Na imagem abaixo há um exemplo de como um arquivo sdf é estruturado.

RDKit	3D
23 26 0 0 0 0 0 0 0 0 0999 V2000	
1.8547 2.2184 1.5929 O 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	
2.0422 1.3139 0.8104 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	
4.1070 -0.8090 -1.3027 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	
3.0756 -0.2045 -0.6132 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	
10 11 1 0	
11 12 2 0	
12 13 1 0	
13 14 2 0	
22 23 2 0	
23 2 1 0	
8 3 1 0	
23 9 1 0	
18 13 1 0	
M CHG 2 19 1 21 -1	
M END	
> <minimizedAffinity>	
-10.02275	
> <minimizedRMSD>	
6.28430	
\$\$\$\$	

No Admetscore, o SDF é o formato de entrada principal. O programa utiliza a biblioteca Selenium para acessar o site do ADMETlab 3.0, realizar o screening e enviar o arquivo SDF com a lista de moléculas para se ter a análise ADMET.

O site possui uma limitação de até 1000 moléculas por vez, mas devido a instabilidades no site o Admetscore envia apenas 300 moléculas por

vez, por padrão, para evitar sobrecarregamento. No entanto, os usuários podem ajustar esse número através de um parâmetro na linha de comando.

Para gerenciar o número de moléculas, o Admetscore divide o SDF de entrada em arquivos menores. Esses arquivos são armazenados em uma pasta chamada "sdf" no diretório onde o programa está executando. Cada arquivo é nomeado com o identificador do banco de dados e o identificador da partição. Ex:

Banco de Dados	Ex: ID de Molécula	Ex: Arquivo de Partição
ChEMBL32	CHEMBL3935244	CHEMBL_0001.sdf
ChemDiv	ChemDiv-0449-0140	ChemDiv_0001.sdf
ChemSpace	CSC101176678	CSC_0001.sdf
Enamine	Z2092511367	Z_0001.sdf
LabNetwork	LN00196149	LN_0001.sdf
MCULE	MCULE-6264776491	MCULE_0001.sdf
MolPort	MolPort-005-944-795	MolPort_0001.sdf
NCI Open Chemical Repository	NSC368697	NSC_0001.sdf
PubChem	PubChem-137222558	PubChem_0001.sdf
ZINC	ZINC000257617615	ZINC_0001.sdf

AdmetLab 3.0

O ADMETlab 3.0 é uma plataforma web projetada para a avaliação de parâmetros ADMET (Absorção, Distribuição, Metabolismo, Excreção e Toxicidade) de moléculas no processo de descoberta de medicamentos. A descoberta de novos medicamentos é um empreendimento complexo, com uma alta taxa de falhas nas fases pré-clínicas e clínicas. O site ajuda a identificar e eliminar compostos com características indesejáveis antes que eles avancem no estado de desenvolvimento

do estudo.

No ADMETlab 3.0 há 119 endpoints diferentes, que são os parâmetros específicos que são avaliados para determinar as propriedades ADMET de uma molécula. Cada endpoint fornece uma medida ou previsão sobre um aspecto particular do comportamento da molécula no organismo. Com isso o site faz uma análise empírica que é a avaliação desses endpoints baseados em dados experimentais e observações de como compostos similares se comportam. Nessa análise há os critérios de aceitabilidade, onde cada endpoint terá seus valores de limites desejáveis, que serão considerados com bons, enquanto valores fora desses limites podem indicar problemas potenciais, considerados como ruins.

O Admetscore usa esses valores de aceitabilidade de cada endpoint para fazer sua classificação, o programa faz sua classificação da seguinte forma: Usa-se uma base decimal onde cada grupo ADMET recebe um total de 10 pontos para ser distribuído entre seus endpoints. Cada endpoint é avaliado e classificado em uma das três categorias: Aceitável, recebe a pontuação máxima atribuída ao endpoint. Médio, recebe metade da pontuação máxima. Ou ruim, não recebe pontuação.

A pontuação de cada endpoint é somada para calcular a nota total do grupo. Então é feita uma média ponderada das notas dos grupos, atribuindo diferentes pesos a cada grupo conforme sua importância na análise geral. Isso assegura que os grupos mais relevantes influenciam proporcionalmente a pontuação final. Os valores dos pesos podem ser alterados ao modificar um arquivo chamado weights.json, que é criado no diretório onde o programa foi executado pelo menos uma vez. Os valores padrão são:

GRUPO	PESO
Absorção	3
Distribuição	1
Toxicidade	8
Tox21	2

GRUPO	PESO
Metabolismo	2
Toxicóforo	3
Excreção	1
Química Medicinal	5

Instalando o Admetscore

1. Baixe o arquivo .whl, o arquivo terá a seguinte nomenclatura, trocando apenas a parte da versão: `admetscore-version-py3-none-any.whl`
2. Para instalar o admetscore é necessário ter instalado o Python com versão igual ou acima de 3.10, e o Pip.
3. No terminal onde o arquivo .whl estiver dê o comando `pip install arquivo.whl` para instalar a ferramenta. Substituindo pelo nome exato do .whl
4. Caso você já tenha instalado alguma versão anterior do admetscore no seu computador é importante que você force a re-instalação com o comando:

```
pip install arquivo.whl --force-reinstall
```

5. Após a instalação, você poderá conferir se ocorreu tudo certo com a instalação visualizando a versão da ferramenta que está instalada no seu computador, com o comando:

```
admetscore -v
```

Utilização Como Ferramenta

```
admet_score [OPÇÕES] -s [Procurar Pelo SDF no Diretório Atual]
```

```
admet_score -i [CAMINHO PARA O SDF] [OPÇÕES]
```

```
admet_score split -i [CAMINHO PARA O SDF] [OPÇÕES]
```

```
admet_score evaluate -i [CAMINHO PARA UM DAS PARTES DO ARQUIVO  
CSV DE RESPOSTA DO SITE VINDO DO ADMETLAB 3.0] -s [CAMINHO  
PARA O SDF CORRESPONDENTE AO CSV] [OPÇÕES]
```

OPÇÕES

-n, --batch_number: *[Não disponível para o módulo evaluate]* Use este parâmetro se quiser alterar o número de moléculas em cada partição. Isso é útil se o site ADMETlab 3.0 aceitar mais ou menos moléculas por screening. O padrão é definido como 300.

-ps, --pharmmit_score_docking: *[Não disponível para o módulo evaluate]* Use este parâmetro para fazer a partição do sdf a partir da pontuação de docking especificada. Isso reduzirá o número de partições criadas, pois, por padrão, todo o arquivo SDF é particionado.

-t, --best_hits_number: *[Não disponível para o módulo split]* Especifique o número de moléculas com pontuação máxima que você deseja visualizar. A planilha não será cumulativa e um arquivo separado será criado na pasta de pontuação. Por

padrão, uma planilha é criada com os 50 melhores resultados.

-s: *[Não disponível para os módulos]* Use este parâmetro para procurar por arquivos sdf no diretório atual. Se o parâmetro for usado, a ferramenta procurará por arquivos sdf no diretório atual e os processará automaticamente.

Utilização Como Biblioteca

```
import admetscore

admetscore.Screening(verbose=bool=False).screen (input_file: str,
batch_size: int = 299, affinity_cutoff: float = None, best_hits_number:
int = 50)
```

```
admetscore.Splitter(verbose=bool=False).split (input_file: str,
batch_size: int = 299, affinity_cutoff: float = None)
```

```
admetscore.Evaluate(verbose=bool=False).evaluate_data (csv_file: str,
sdf_file: str, best_hits_number: int)
```

Criando Distribuição Binária

Para instalar a ferramenta, é necessário um arquivo .whl (Wheel), que permite o uso do pip para a instalação. O código do admetscore está disponível no repositório do GitHub. O empacotamento do projeto foi feito utilizando o Poetry para gerenciar as dependências e a configuração, e o Python Build para criar o arquivo

.whl. O arquivo pyproject.toml já está configurado no código. Para gerar o arquivo .whl, siga os passos abaixo:

- Instale as bibliotecas necessárias:

- Para instalar o Poetry:

```
pip install poetry
```

- Para instalar o Python Build:

```
pip install build
```

-

- Crie um arquivo chamado pyproject.toml e coloque o seguinte texto:

```
[tool.poetry]

name = "admet_score"

version = "1.0.0"

description = "The ADMETSCORE is a software that aims to facilitate the ADMET analysis of molecules resulting from pharmacophore research."

authors = ["Julio Cesar Xavier <jcaxavier2@gmail.com>"]

readme = "README.md"

include = ["weights.json"]

[tool.poetry.dependencies]

python = "^3.10"

pandas = "^2.2.2"

openpyxl = "^3.1.5"

rdkit = "^2024.3.5"

selenium = "^4.23.1"

webdriver-manager = "^4.0.2"
```



```
XlsxWriter = "^3.2.0"

numpy = "1.26.4"

[build-system]

requires = ["poetry-core"]

build-backend = "poetry.core.masonry.api"

[tool.poetry.scripts]

admet_score = "admet_score.app:main"
```

- Após configurar o `pyproject.toml`, caso as bibliotecas exigidas pela ferramenta ainda não estejam instaladas no ambiente de código, execute o seguinte comando para instalar todas as dependências listadas no arquivo `pyproject.toml`:

```
poetry install
```

- Com o arquivo `pyproject.toml` pronto e as dependências instaladas, crie o arquivo `.whl` com o seguinte comando:

```
python -m build
```

- O arquivo `.whl` será gerado na pasta `dist/` do projeto.

Após a criação do arquivo Wheel, ele estará pronto para ser distribuído e instalado via `pip`.