split\_data.py contains the data splitting method, extracting features from the dataset using getdata(), obtain motif, and calculate molecular graph features to derive two feature matrices. It then computes three types of molecular fingerprint features using the rdkit toolkit. utils.py includes functions used for dataset split. layer.py incorporates a graph attention network. The model.py file encompasses the basic architecture of the model. In this file, the feature extraction of the three molecular representations is completed and fused together, return the model output after apply the sigmoid activation function. train\_evaluate.py includes train and evaluation processes, implement 10-fold cross-validation in the outermost loop, with the model hyperparameters located in this file. Follow the model's prediction results, performance evaluation of the model is conducted within this file. The train and evaluation results are stored in the results folder.

# AMPred-MFG: Investigating the mutagenicity of compounds using motif-based graph combined with molecular fingerprints and graph attention mechanism

## fingerprint ECFP、MACCS、RDkit can be extracted by rdkit

\*\*\*

# Environment：

absl-py 2.1.0

alabaster 0.7.13

annotated-types 0.6.0

autopep8 2.1.0

Babel 2.14.0

Brotli 1.0.9

cachetools 5.3.3

certifi 2024.2.2

charset-normalizer 2.0.4

click 8.1.7

cloudpickle 3.0.0

colorama 0.4.6

contourpy 1.1.1

cycler 0.12.1

dgl-cuda11.3 0.9.1

dglgo 0.0.2

dgllife 0.3.2

docutils 0.20.1

filelock 3.13.1

fonttools 4.51.0

future 1.0.0

gmpy2 2.1.2

google-auth 2.29.0

google-auth-oauthlib 1.0.0

grpcio 1.62.2

hyperopt 0.2.7

idna 3.4

imagesize 1.4.1

importlib\_metadata 7.1.0

importlib\_resources 6.4.0

isort 5.13.2

Jinja2 3.1.3

joblib 1.4.0

kiwisolver 1.4.5

littleutils 0.2.2

Markdown 3.6

markdown-it-py 3.0.0

MarkupSafe 2.1.3

matplotlib 3.7.5

mdurl 0.1.2

mkl-fft 1.3.8

mkl-random 1.2.4

mkl-service 2.4.0

mpmath 1.3.0

networkx 3.1

numpy 1.24.3

numpydoc 1.7.0

oauthlib 3.2.2

ogb 1.3.6

outdated 0.2.2

packaging 23.2

pandas 2.0.3

pillow 10.2.0

pip 23.3.1

platformdirs 3.10.0

pooch 1.7.0

protobuf 5.26.1

psutil 5.9.0

py4j 0.10.9.7

pyasn1 0.6.0

pyasn1\_modules 0.4.0

pycodestyle 2.11.1

pydantic 2.6.4

pydantic\_core 2.16.3

Pygments 2.17.2

pyparsing 3.1.2

PySocks 1.7.1

python-dateutil 2.9.0.post0

pytz 2024.1

PyYAML 6.0.1

rdkit 2023.9.5

rdkit-pypi 2022.9.5

requests 2.31.0

requests-oauthlib 2.0.0

rich 13.7.1

rsa 4.9

ruamel.yaml 0.18.6

ruamel.yaml.clib 0.2.8

scikit-learn 1.3.2

scipy 1.10.1

seaborn 0.13.2

setuptools 68.2.2

shellingham 1.5.4

six 1.16.0

snowballstemmer 2.2.0

Sphinx 7.1.2

sphinxcontrib-applehelp 1.0.4

sphinxcontrib-devhelp 1.0.2

sphinxcontrib-htmlhelp 2.0.1

sphinxcontrib-jsmath 1.0.1

sphinxcontrib-qthelp 1.0.3

sphinxcontrib-serializinghtml 1.1.5

sympy 1.12

tabulate 0.9.0

tensorboard 2.14.0

tensorboard-data-server 0.7.2

threadpoolctl 3.4.0

tomli 2.0.1

torch 2.0.0

torchaudio 2.0.0

torchdata 0.7.1

torchvision 0.15.0

tqdm 4.65.0

typer 0.12.2

typing\_extensions 4.9.0

tzdata 2024.1

urllib3 2.1.0

Werkzeug 3.0.2

wheel 0.41.2

win-inet-pton 1.1.0

zipp 3.18.1