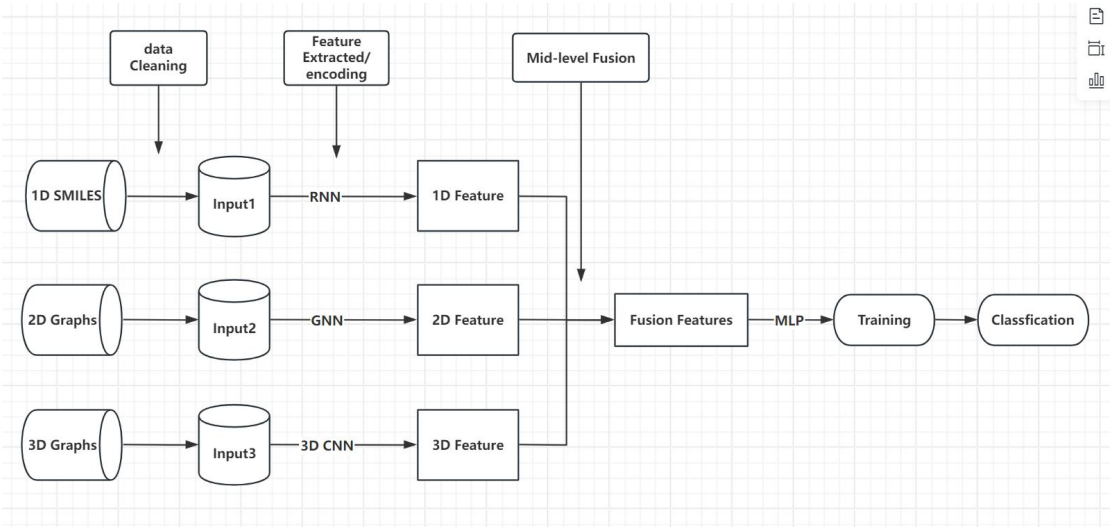
Integrate multimodal molecular data to predict the properties of molecular compounds.
The general structure of the **Baseline** is as follows.



1, Get 1D, 2D, 3D data as input;
2, Model the data of each mode separately and extract the features (feature encoding);
3, merged the three features into a unified feature set, that is, a high-dimensional vector.
4, Use deep learning model(like MLP) to train the combined feature set.

More Details:
Data:
1, As for the original data,they can be found on data websites;
I am trying to find the data;
There are two strategies:
(1) Obtain data of different dimensions of the same molecular compounds directly;
(2) Use professional software such as RDkit to transform low-dimensional data into high-dimensional data;
2, May do some data augments to increase the diversity of data.

Feature Extracted/encoding:
1, RNN for 1D SMILES data
2, GNN(like ML-MPNN) for 2D graphs data;
3, 3D-CNN/Point cloud processing algorithms (such as PointNet) extract spatial structure features.

Feature fusion:
1, By means of Mid-level Fusion, data feature vectors of different modes are merged into a single high-dimensional feature vector by means of concatenation or weighted average.
2, The multimodal fusion section has many aspects worthy of experiment and research, such as the stage of fusion (early, middle, late), and the way of fusion.
3, After feature fusion, normalization may be performed to ensure consistent proportions between different features

Training:

1, When training integrated features, consider using a deep learning model first:
for example, a multi-layer perceptron (MLP)

2, but also fully connected Neural Networks (Dense Neural Networks) and support vector machines (SVM) or Random Forests.