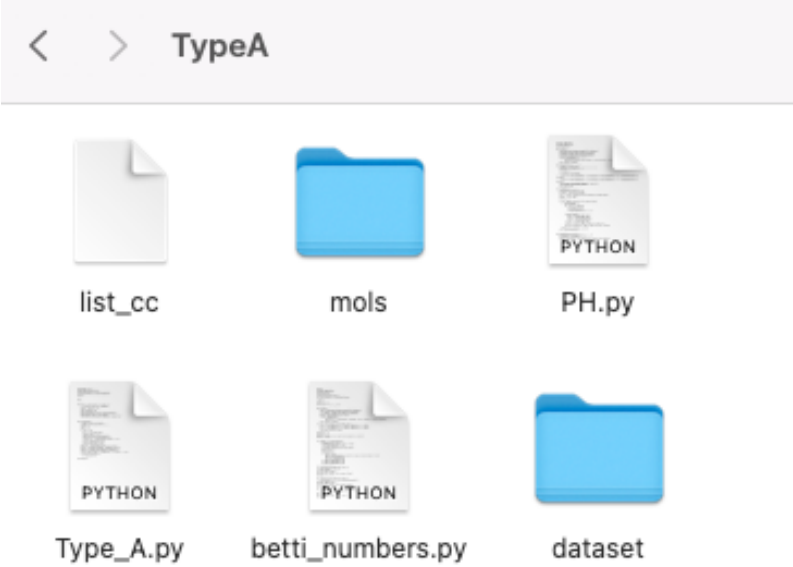
README file for Type A



1. list_cc : the dataset from paper “Computational Predictions of Glass-Forming Ability and Crystallization Tendency of Drug Molecules”.

2. mols: the conformation files of each molecule in the dataset are stored in XYZ format.

3. PH.py: the relevant algorithms for persistent homology.

4. Type_A.py: the code to generate the feature matrix using Type A descriptors.

```python
import pandas as pd
from PH import CombinedFeature
import collections
collections.Iterable = collections.abc.Iterable
from sklearn.preprocessing import MinMaxScaler
import os

a = 1
b = 3

n = 10
```

a, b, n are consistent with those in the article. a and b are fixed. You can modify n to obtain descriptors with different levels of discretization of the filtering variable ϵ, and n also determines the number of predictors, which is given by the formula $3n + 15$.

5. betti_numbers.py: the code to plot Figure S1 in the SI, which is the reason why we set $a = 1$ and $b = 3$.

6. dataset: the feature matrices corresponding to different values of n are used for model selection.