**Datasets and codes of Drug R+**

In addition to python code, this repository includes the created datasets. For every dataset, two txt files are available in it:

1- Dataset: this file contais known drug-target.

2- Potentials\_drugs\_and\_targets: This file entails all potential drug-target interactions (unknown interactions) exclude known drug-target interactions.

These files consistis of:

1- drug: name of drug such as D.

2- target: name of target (enzyme) such as T.

3- S1: number of drugs which interact T exclude D.

4- S2: average similarity score between D and drugs of T.

5- S3: minimum similarity score between D and drugs of T.

6- S4: maximum similarity score between D and drugs of T.

7- S1: number of targets of T exclude D.

8- S2: average similarity score between T and targets of D.

9- S3: minimum similarity score between T and targets of D.

10- S4: maximum similarity score between T and targets of D.

The python codes and script of the datasets are gathered in the following files:

1. DrugR: Scripts of the database (DDL)
2. Repurpose: A store procedure for finding drugs based on their MOA and common targets.
3. Map: A python code for dividing the flat file into small files.
4. Reduce: A python code for traversing the files and transferring them into the database.