ReadMe:

1. Final report is submitted as “final report.pdf”
2. Presentation slide is submitted as “presentation.pdf’
3. Code folder
   1. MIMIC-3 knowledge and Drug-Drug Interaction extraction: followed by the code provided from <https://github.com/ycq091044/SafeDrug.git>
   2. Processing the output file from 3.1 to unduplicated knowledge graph, is submitted as “KG\_processing.ipynb”
   3. DiffKG model implementation: the main train and test loop are defined and submitted as “mediff.py”
4. Data
   1. Rawdata: MIMIC3 dataset from <https://physionet.org/content/mimiciii/1.4/> (huge dataset, I didn’t submit it)

NDC-RXCUI-ATC4 file from <https://github.com/sjy1203/GAMENet>,

CID-ATC file from <https://github.com/sjy1203/GAMENet>,

NDC-RXCUI file from <https://github.com/sjy1203/GAMENet>,

Drug back file from <https://drive.google.com/file/d/1EzIlVeiIR6LFtrBnhzAth4fJt6H_ljxk/view?usp=sharing>,

Drug-DDI file from <https://drive.google.com/file/d/1mnPc0O0ztz0fkv3HF-dpmBb8PLWsEoDz/view?usp=sharing>

* 1. Processed knowledge graphs
     1. MIMIC3 knowledge graphs: submitted as “mimic3\_records.pkl” and “vocabulary\_mapping.pkl”
     2. Drug interaction knowledge graph: submitted as “ddi.pkl”
  2. knowledge graph

submitted as splitted files: “kg.txt”, “trnMat.pkl” and “tstMat.pkl”