GCMC R&D

막히는 부분

1. Docker에서 컨테이너를 통해 분석환경을 구현할 때, GPU를 사용 못함 (엔비디아 도커 설치 및 적용 필요)
2. 구현된 코드에서 데이터 셋을 전처리 되어 있는 데이터 셋을 사용하므로 실 데이터에 적용하기 위해서 데이터셋을 불러오는 함수를 고쳐야 한다.
3. PyG의 데이터셋 customized랑 loader부분에 대한 이해가 부족하여 실 데이터를 적용하는 함수를 구현하는데 한계를 느낌
4. 연산에서 자료열 구조가 일치하지 않는 문제

해결

1. GPU 사용 없이 데이터셋의 일부를 잘라 서브그래프 느낌으로 코드 구현 및 실행
2. 개인의 csv 파일을 압축 및 url화 시켜서 해당 os에 다운로드 하고 path를 지정하여 불러오는 방법으로 해결
3. pyg 데이터, 데이터셋의 함수의 input 속성 및 사용방법 예시를 학습하여 기존의 csv파일을 해당환경에 불러오고 파이토치 형식의 Data로 직접 만들어 dataset으로 사용하는 방식으로 해결
4. 데이터 프레임 혹은 변수 지정에서 int와 str이 같이들어가서 부딪치는게 없는지 코드를 일부씩 뜯어 확인하는 방법으로 해결