[현재 상황]

농사로에 있는 식물 데이터와 플립 사이트에 있는 식물 관련 데이터 추출 완료하였습니다.

추출한 데이터와 관련한 데이터 분석 진행 중에 있으며, 추후 협업 필터링을 위해 오늘의 집과 심폴에 있는 리뷰 데이터를 크롤링하고 있는 중에 있습니다.

[향후 계획]

농사로와 플립 사이트에 있는 식물 데이터에서 특성을 추출하여 컨텐츠 기반 필터링을 진행할 예정입니다.

오늘의 집과 심폴 리뷰 데이터를 활용하여 협업 필터링을 해볼 생각입니다.

[QnA]

크롤링한 리뷰 데이터를 활용하려고 할 때, 분산되어 있는 사용자 정보는 협업 필터링을 사용하는데 있어서 괜찮은 데이터인가요?

괜찮지 않은 데이터면은 약간은 저희가 보안 작업을 진행해도 되는 건가요?

컨텐츠 기반 필터링을 유사도를 활용하여 추천을 진행할 것인데 주의해야 할 것이 있나요?

데이터 베이스에 저장을 할 때, 컬럼 값을 코드로 주어서 참조테이블을 하는 것이 좋나요?

아니면은 그냥 값을 저장하는 것이 좋을까요?