**Weekly report**

날짜 : **2023-08-18**

연구원 : **최윤석**

* **이전 수행계획**
  + **관계형 테이블 임베딩 :**
    - CNE-join의 성능을 개선할 방법 고민
      * DIALITE: Discover, Align and Integrate Open Data Tables가 참고한 논문들을 읽어보면서 새로운 아이디어 캐치
    - 박찬훈 연구원과 DataMap 개발(Type A, B)
      * A. DataMap에서 Top-N개의 조인가능한 테이블 목록 반환하는 코드 작성
      * A. DataMap에서 조인가능한 테이블을 자동으로 조인하는 코드 작성
      * A. DataMap을 그리기 위해 테이블간 유사도 계산
      * B. 질의 방식(Query Table을 입력으로 주는 방식)의 CNE-join 만들기

* **수행결과**
  + **관계형 테이블 임베딩 :**
    - 박찬훈 연구원과 DataMap 개발
      * A. DataMap에서 Top-N개의 조인가능한 테이블 목록을 반환하는 코드 작성 완료
        + 조인 가능한 두 테이블명과 조인 컬럼, 그리고 두 테이블간의 조인 가능성을 0에서 1사이의 값으로 나타내고 조인 가능성을 기준으로 내림차순으로 제시

텍스트, 흑백, 폰트, 종이이(가) 표시된 사진

자동 생성된 설명

* + - * A. DataMap에서 조인가능한 테이블을 자동으로 조인하는 코드 작성 완료
        + 두 테이블과 조인 컬럼이 주어지면 두 테이블을 조인한 결과를 반환
      * A. DataMap을 그리기 위해 테이블간 유사도 계산하는 코드 작성 완료
        + 테이블의 모든 컬럼명을 FastText를 이용해 300차원의 벡터로 바꾸고, 컬럼명 벡터를 모두 더해 평균을 낸 벡터를 테이블의 대표 벡터로 지정함
        + 두 테이블 벡터간 코사인 유사도 계산을 통해 두 테이블의 유사도를 알아냄

텍스트, 스크린샷, 폰트, 흑백이(가) 표시된 사진

자동 생성된 설명

* **수행계획**
  + **관계형 테이블 임베딩 :**
    - CNE-join의 성능을 개선할 방법 고민
      * DIALITE: Discover, Align and Integrate Open Data Tables가 참고한 논문들을 읽어보면서 새로운 아이디어 캐치
    - 박찬훈 연구원과 DataMap 개발(Type A, B)
      * A. Top-N개의 조인 가능한 테이블을 반환하는 방식의 문제점(조인 가능성 값이 높은데 실제 조인 가능하지 않은 케이스가 존재)을 보완
      * B. 질의 방식(Query Table을 입력으로 주는 방식)의 CNE-join 만들기
* **기타사항**