**Weekly report**

날짜 : **2023-12-13**

연구원 : **최윤석**

* **이전 수행계획**
  + **관계형 테이블 임베딩 :**
    - CNE-join의 성능을 개선할 방법 고민
      * 조인하는 테이블의 개수를 3개 이상으로 확장해보기, 최적의 테이블 개수는 무엇인가? 이를 확인할 척도에 대해서도 고민
        + CNE-join의 결과로 나온 2개의 조인 가능한 테이블쌍 중에서 A-B, B-C 같이 테이블이 겹치는 쌍을 활용해 n-join-pair를 생성하는 알고리즘 및 이에 대한 융합도 산정 방안 고민 및 소스코드 작성(이전 CNE-join과의 일관성이 중요함)
      * 시간 단축 문제는 가장 나중으로 진행
    - Measuring and Predicting the Quality of a Join for Data Discovery 논문이 참고한 문헌을 읽고 n-join-pair 알고리즘 아이디어 얻기
    - CNE-join 결과로 나온 조인 가능한 테이블쌍 정보를 활용해 최대 4개의 테이블을 조인한 테이블을 벡터로 임베딩
      * 조인한 테이블의 임베딩 벡터와 조인에 사용된 테이블의 임베딩 벡터간의 거리를 활용해 n-join-pair에 관한 insight(Tier제 융합결과 랭킹법)를 얻기
    - 박찬훈 연구원과 DataMap 개발(Type A, B)
* **수행결과**
  + **관계형 테이블 임베딩 :**
    - 조인한 테이블의 임베딩 벡터와 조인에 사용된 테이블의 임베딩 벡터간의 거리를 활용해 n-join-pair에 관한 insight(Tier제 융합결과 랭킹법)를 얻기
      * 컴퓨터가 다시 시작되어서 새로 임베딩 벡터를 만드는 중, 이후에는 임베딩 벡터를 저장하여 사용
    - Measuring and Predicting the Quality of a Join for Data Discovery 논문이 참고한 문헌을 읽고 n-join-pair 알고리즘 아이디어 얻기
      * BigGorilla: An Open-Source Ecosystem for Data Preparation and Integration 논문 읽음
        + 2개의 테이블 데이터가 주어졌을 때, 2개의 테이블을 조인한 테이블을 출력해주는 프로그램인 BigGorilla에 대해 소개하는 논문
        + n-join-pair와는 큰 관련이 없음
  + **빅데이터 자료구조 조교 :**
    - 빅데이터 자료구조 모든 과제 채점 및 엑셀에 과제 성적 기입 완료
* **수행계획**
  + **관계형 테이블 임베딩 :**
    - CNE-join의 성능을 개선할 방법 고민
      * 조인하는 테이블의 개수를 3개 이상으로 확장해보기, 최적의 테이블 개수는 무엇인가? 이를 확인할 척도에 대해서도 고민
        + CNE-join의 결과로 나온 2개의 조인 가능한 테이블쌍 중에서 A-B, B-C 같이 테이블이 겹치는 쌍을 활용해 n-join-pair를 생성하는 알고리즘 및 이에 대한 융합도 산정 방안 고민 및 소스코드 작성(이전 CNE-join과의 일관성이 중요함)
      * 시간 단축 문제는 가장 나중으로 진행
    - Measuring and Predicting the Quality of a Join for Data Discovery 논문이 참고한 문헌을 읽고 n-join-pair 알고리즘 아이디어 얻기
    - CNE-join 결과로 나온 조인 가능한 테이블쌍 정보를 활용해 최대 4개의 테이블을 조인한 테이블을 벡터로 임베딩
      * 조인한 테이블의 임베딩 벡터와 조인에 사용된 테이블의 임베딩 벡터간의 거리를 활용해 n-join-pair에 관한 insight(Tier제 융합결과 랭킹법)를 얻기
    - 박찬훈 연구원과 DataMap 개발(Type A, B)
    - 자동 융합 시나리오에 대한 명세를 작성해보기(use case diagram)
      * PPT로 작성
    - 융합을 할 때 Join을 할 지 Union을 할 지 확인하는 방법 – 나는 Join 파트
* **기타사항**