**Weekly report**

날짜 : **2023-10-18**

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
    - CNE-join 특허출원 관련 최항구 변리사님과 미팅(10/17 화요일 오후 2시 화상미팅)
    - 박찬훈 연구원과 DataMap 개발(Type A, B)

* **수행결과**
  + **관계형 테이블 임베딩 :**
* n-join-pair를 생성하는 알고리즘 구상완료 및 이를 구현하는 코드 작성 중
* n-join-pair 생성 알고리즘 소개

스케치, 도표, 그림, 원이(가) 표시된 사진

자동 생성된 설명

* 1. CNE-join 결과로 나온 조인 가능한 테이블 쌍을 활용해 트리 자료구조를 생성
  + 조인 가능한 테이블쌍에 나온 모든 테이블을 한 번씩 루트 노드로 하여 위의 그림같이 여러 개의 트리를 만듦
  + 조인 가능한 테이블쌍에 나온 테이블 1개당 하나의 트리가 만들어짐. N개의 테이블이 나왔다면 N개의 트리를 만듦
* 2. 각 트리에서 나올 수 있는 모든 조인 가능한 경우의 수를 구함
  + 예시)   
    테이블 A가 루트 노드인 트리의 경우 A-B-C / A-B-D / A-C-D / A-B-C-D / A-B-E / …  
    테이블 B가 루트 노드인 트리의 경우 B-A-E / B-A-F / B-E-F / B-A-C / …
* 3. 각각의 경우에 대해 아래의 식을 만족하는지 확인한다
  + 아래의 식을 만족하면 n-join-pair에 부합하다고 판단할 것임
  + > Threshold
  + 식에 대한 근거가 필요함(예시나 원칙 규칙 등 논거가 필요해 보임)
  + Inner Join 만이 아닌 Outer Join같은 다른 조인 방법도 고려해봐야 함
  + 식을 다른 방면으로도 생각해서 개량해볼 생각도 해볼 것

텍스트, 스크린샷, 도표, 라인이(가) 표시된 사진

자동 생성된 설명

* n-join-pair를 생성하는 코드를 짜면서 생각난 문제점
  + 위와 같은 경우에 nobel\_latest.csv와 Food\_Preference.csv 테이블을 조인을 해야 할지 하지 말아야 할지 잘 모르겠음
* CNE-join 특허출원 관련 최항구 변리사님과 미팅 진행
  + 변리사님께서 CNE-join 작동원리에 대해 물어보시고 이에 대해 답변하는 방식으로 미팅을 진행함
* **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
    - 박찬훈 연구원과 DataMap 개발(Type A, B)
* **기타사항**