**Weekly report**

날짜 : **2023-01-12**

연구원 : **최윤석**

* **이전 수행계획**
  + **차량 이미지 분류** :
    - 1차 보고서 내용 수정
    - SSD 소스코드 파일 정리 및 소스코드에 대한 간략한 설명 작성
  + **관계형 테이블 임베딩** :
    - 수치형 데이터를 이산화한 것을 범주형 데이터로 바꾸는 임베딩 방식을 보완하는 방법을 찾고 실험하여, 기존의 윤종찬 연구원의 연구와 **융합 가능도 척도**를 사용하여 성능 비교
    - 각 컬럼마다의 중요도를 고려해서 임베딩하는 방식과 수치형 컬럼에서 숫자의 크기가 크고 작음에 부여된 의미를 고려하는 임베딩 방식 연구
    - 나만의 임베딩 아키텍처 / 융합 가능성 아키텍처(임베딩 아키텍처에 dependant한) 그려보기, 기존의 것과 나만의 것을 분리하여 정리 (PPT로)
    - 모델의 성능이 안 좋은 이유 찾고 해결. (데이터의 문제 / 임베딩 방식의 문제) – 데이터의 문제 : 융합 가능한 라벨들을 찾아서 추가 , 임베딩 방식의 문제 : 이전 논문에서 읽었던 Field-wise Network 방식을 적용
  + **GAN 논문 세미나 :**
    - StackGAN 논문 읽고 세미나 준비 (PPT 30분 이내 발표)
* **수행결과**
* **차량 이미지 분류 :**
* 차량 이미지 분류 보고서 수정사항 수정하여 최종보고서 작성
* 차량 이미지 분류 소스코드 정리 및 각 소스코드에 대한 설명파일 작성
* SSD 소스코드 파일 박종명 부장님과 공유
* **관계형 테이블 임베딩** : (임베딩 / 융합 가능성 파트 분류해서 작성)
* **GAN 논문 세미나 :**
  + StackGAN 논문 읽고, 발표 PPT 작성
* **수행계획**
  + **차량 이미지 분류** :
  + **관계형 테이블 임베딩 :**
    - 수치형 데이터를 이산화한 것을 범주형 데이터로 바꾸는 임베딩 방식을 보완하는 방법을 찾고 실험하여, 기존의 윤종찬 연구원의 연구와 **융합 가능도 척도**를 사용하여 성능 비교
    - 각 컬럼마다의 중요도를 고려해서 임베딩하는 방식과 수치형 컬럼에서 숫자의 크기가 크고 작음에 부여된 의미를 고려하는 임베딩 방식 연구
    - 나만의 임베딩 아키텍처 / 융합 가능성 아키텍처(임베딩 아키텍처에 dependant한) 그려보기, 기존의 것과 나만의 것을 분리하여 정리 (PPT로)
    - 모델의 성능이 안 좋은 이유 찾고 해결. (데이터의 문제 / 임베딩 방식의 문제) – 데이터의 문제 : 융합 가능한 라벨들을 찾아서 추가 , 임베딩 방식의 문제 : 이전 논문에서 읽었던 Field-wise Network 방식을 적용
    - 연구재단 보고서 작성(1/18까지 교수님께 메일로) 최대한 빨리
    - 데이터셋 융합 관련 논문 2편 읽고 공부
  + **GAN 논문 세미나 :**
    - **Towards the Automatic Anime Characters Creation with Generative Adversarial Networks** 논문 읽고 세미나 준비 (PPT 30분 이내 발표) , 각 논문의 핵심사항을 알아내는게 목적. 연구 동기, 배경, 관련연구 등은 생략하고 이해한 범위 내에서 핵심사항 중심으로 발표 진행. 성능 평가는 한 페이지(메인 그래프만)로 충분함. 그림과 키워드 위주로 발표.
* **기타사항**
  + **데이터 융합 프로세스**



* **융합 가능도 척도**
  + 임베딩 거리 기준 값을 정하고 기준값보다 임베딩 거리가 가까우면 융합 가능한 데이터셋이라 판단함
  + 융합 가능한 데이터셋을 사람이 하나씩 확인하여 유의미한 융합이 되는 데이터셋을 찾음
  + 을 융합 가능도 척도라 정의함



혼합형 데이터셋 예시

* **1월 초 연구 내용 정리 후 발표**
* **StackGAN 피드백 : 정통적인 GAN과의 차이점의 부재, 페이퍼가 주는 시사점 파악, 성능평가는 좀 더 미니멀하게 진행, 목차는 필요없음, 데이터셋은 뭘 썼는지 앞에 설명 필요, 각 변수가 등장하면 무슨 뜻인지 설명 필요**