**Weekly report**

날짜 : 2022-06-23

연구원 : 최윤석

* 수행결과

▶ 윤종찬 연구원의 석사논문 열람

* 수행계획

▶ 한국전자거래학회 발표문 및 PPT 제작과 리허설 영상 찍기

▶ 윤종찬 연구원의 실험 코드를 받아 관계형 테이블 임베딩 방식 이해(코드 돌려보면서 확인, 입력 데이터와 출력 데이터 이해, 7월 중순부터 8월까지는 새로운 연구 아이디어를 찾아야 함. 업그레이드 or 새로운 방향)

▶ A survey on Image Data Augmentation for Deep Learning 자료 읽기 ( 트라콤 )

▶ inference 모델이 로컬환경에서 돌아갈 때 발생하는 오류 해결

▶ 오류를 해결하더라도, 10만개의 뉴스데이터를 돌리려면 500시간 이상이 걸릴 것으로 예상, 1만개의 뉴스데이터도 50시간 이상으로 예상됨. 따라서 수행시간의 대부분을 차지하는 inference 코드를 빠르게 돌릴 방법이 없는지 확인

▶ 고유명사와 개체명 인식할 때, 단어에 조사가 붙는 경우가 많은데 (ex: 최윤석은 – NNP – PER) 이를 해결할 방안 찾기

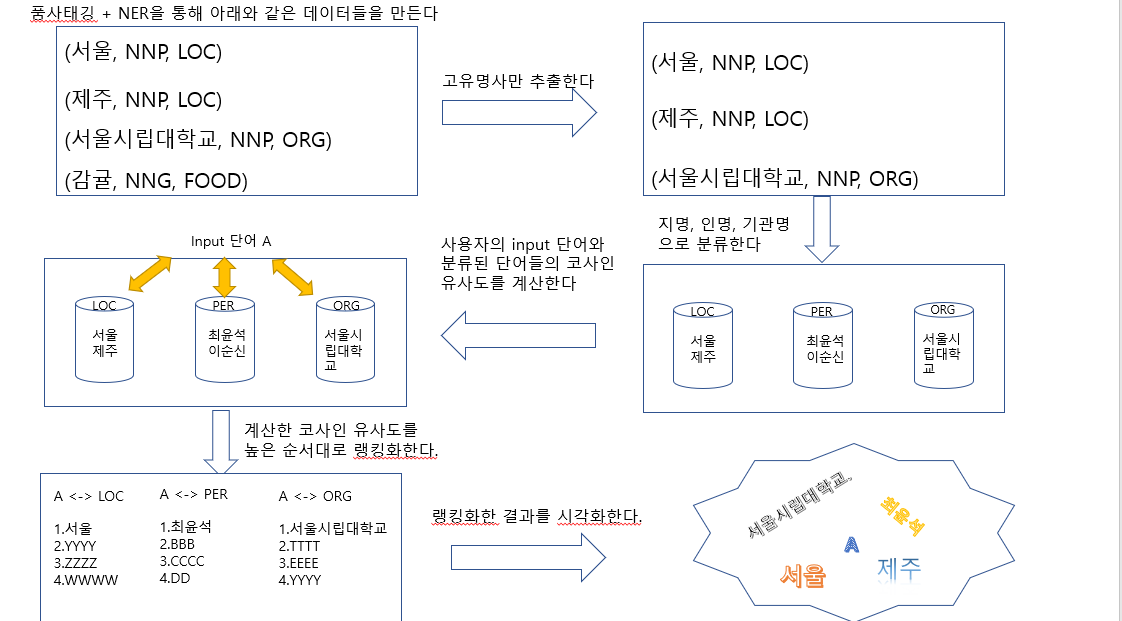


Figure 2. 앞으로의 진행 흐름

* 기타사항

▶ 그림 1, 2에서 LOC(지명) 식으로 설명 추가

▶ 그림 1 사람그림 인터넷에서 괜찮은 아이콘으로 하나 구해서 쓰기

▶ 그림 2 품사태깅 + NER 파트등을 모두가 이해할 수 있게 일반적인 단어로

▶ 그림 2 프로세스 순서 화살표에 1. 고유명사만 추출2,3,4 넣는식으로

▶ 중요한 키워드를 나열하는 개조식으로 문장 쓰기.

▶ A 대신 Q 사용

▶ LOC PER ORG 이외에도 그림 2에 더 추가할 수 있음을 알려주는 것 추가

▶ 그림 3 고유명사를 입력해주세요 빼기, 고유명사 -> 입력단어, 질의단어로 수정

▶ 그림 3 그림 4는 이미지 편집을 통해 한 페이지에 나오는 것처럼 표시

▶ 그림 4 이미지 편집을 통해 가로로 나오게

▶ 예시 단어 손흥민 , 코로나, 등등 일반적인 것으로.