**Weekly report**

날짜 : 2021-10-21

연구원 : 최윤석

* 수행결과

1. 김민중 선배 연구원과 NER을 위한 방향 수립

* 수행계획

1. Entity2vec 코드 이해하고 돌려보기 / 관련 논문인 개체 링킹을 위한 RDF 지식그래프 기반의 포괄적 상호의존성 짝 연결 접근법 논문 스터디
2. NER 을 하기 위한 모델 중 하나인 Conditional Random Field 모델 학습

* 기타사항

1. Entity 안에 Named Entity(고유명사) – 사람이름 , 기관, 위치 등을 문장에서 알아내는게 NER이다. 모든 고유명사를 타겟하기 보다는 사람만 / 기관만 / 위치만 같이 하나만 타겟팅 하는게 좋다.
2. 고유명사가 서로 연관되어 있으면 벡터스페이스에서 서로 가깝게 임베딩 되어야한다.
3. 이후에는 사람 기관 위치를 다 엮어서 연관된 것을 알 수 있다.
4. NER은 고유명사만의 특화된 정보를 담아야 한다. Word2Vec는 이러한 특수성을 잘 담지 못한다. 고유명사만의 특수성은 여러 문장을 통해 나타날 수 있는데, Word2Vec은 여러 문장을 통해 정보를 얻기 힘들다.
5. <https://arxiv.org/abs/1909.01065>
6. https://www.researchgate.net/publication/297752435\_Named\_Entity\_Recognition\_using\_Word\_Embedding\_as\_a\_Feature