**Weekly report**

날짜 : 2021-10-14

연구원 : 최윤석

* 수행결과

1. Contextualized Embeddings in Named-Entity Recognition : An Empirical Study on Generalization 및 Modeling Named Entity Embedding Distribution into Hypersphere 논문 공부 및 이와 관련해 김민중 연구원과 연구방향 설정

* 수행계획

1. ELMO/BERT 와 Fast text 두가지 모델을 같이 사용한 선행 연구논문에 대해 학습(미리 진행된 연구가 있는지 혹은 이게 실현 가능한 것인지. ELMO(문장수준) Fasttext(단어수준) 또한 BERT의 서브워드 토크나이져 만으로 이와 비슷한 수준에 이를 수 있는지)
2. ELMO/BERT 단독으로 NER에 사용하는 것 / 파인튜닝으로 특정 도메인에 대응할 수 있는지에 대해 학습
3. 고유명사 임베딩 (인접성을 형태소 분석같은 방법을 통해 알아낸다 등의 방법을 통해)

* 기타사항

1. ELMO/BERT 두 가지 LM은 성능은 비슷하게 좋았지만 논문에서 ELMO가 다양한 도메인에서 안정된 임베딩 성능을 보여준다고 해서 ELMO + Fasttext를 생각해 봤습니다. (하지만 대부분 BERT를 사용하는 것 같아서 BERT를 중점으로 학습해 볼 생각입니다)