**#8월 19일 스마트타이틀 주간회의**

**(강남 유오워크 / 전원 참석 )**

**<고려대 이기창 전문가 미팅 결과>**

안암에서 이기창님 뵙고 왔습니다. 이기창님은 서울대 국문과를 졸업하시고 6년간 매일경제신문에서 기자로 근무하신 뒤, 현재 고려대학교 산업공학과에서 text mining, NLP, ML등으로 석사과정 공부하시는 분입니다. 연구 관련해서 리서치 도중 관련 포스트를 보고 감명받아 한번 찾아뵈야겠다는 생각이 들어 만나 뵙고 인사 드렸습니다. 이 분은 프로그래밍을 실재로 공부한 것은 1년남짓으로 얼마 되지 않지만, 본인 분야에서의 포스는 대단했고 많이 놀랐습니다.(블로그 애용하게되었습니다) 현재 word-embadding 쪽으로 연구중 이라고 합니다.

본론으로 들어가, 기창님께서는 Seq-to-Seq를 이용해 headline generator 만드는 것은 쉬운일은 아닐 것이라 하셨습니다. 현재 본인이 알고 있는 deep learning 기술로 완벽한 문장을 creative하게 만드는 것은 쉽지 않다 하셨습니다. 때문에 문제를 더 쉬운 방향으로 선회할 것을 권유하셨고 딥 러닝 외의 방법도 고려해 보라 권유하셨습니다. 구체적으로는 첫 문단에 주로 중요한 내용들이 밀집하였으니 그것을 활용해 보라는 말씀, 중요 문장만을 추출하여 이를 이용해 패러프레이징하는 등의 방법을 이용해 보는 것도 말씀해 주셨습니다.

저희 중간과제와 관련해서는 블로그에 완벽하게 포스팅해 주셨습니다. 상세한 설명과 코드가 한국어로 나와있으니 그에 관해 구현하는 방향으로 가면 좋을 것 같습니다.

<https://ratsgo.github.io/natural%20language%20processing/2017/03/19/CNN/>

CNN을 이용해 구현한 방법인데 인상적 이였습니다.

1. **회의본론**

이기창님 미팅이 상당히 의미가 있었다. 특히 첫 문단의 중요성을 강조하면서 방법을 강구하는 것이 중요하다는 조언을 받아들였으면 좋겠다. 첫 번째 문단의 경우는 문장을 뽑아내는 방법론에서 RNN을 어떻게 학습시킬지가 관건일 것 같다. 추천 받은 Stanford 논문을 중심에 두고 최종과제의 결과물을 만들어 가는 것이 좋겠다.

(방향) word-embadding은 벡터를 얻기 위한 목적이 있으며 워드투백터 부분의 공부를 먼저 시작하는 것이 좋을 것 같다. 그리고 CNN공부도 병행하는 것이 좋겠다. 논문이나 포스팅의 샘플을 보면서 스터디를 하고 github에 커밋도 하면 효율적인 프로세스가 가능할 것으로 봄.

(기철) post-by-post로 진행 방향을 먼저 한번 짜보겠다.

NLTK의 깊이가 더 있기도 하고 영문버전을 메인으로 개발을 하며 중간평가를 준비하는 것이 좋겠다. 중간평가 이후에 한국어 버전에 대한 개발방안을 논의해보자.

(충훈) Github관련된 부분을 조금 더 준비해 보겠다.

(기훈) 논문 관련 정리에 도움을 주겠다.

(

**5. 기타 회의**

- 프로젝트 관리의 프로세스를 적용 > 보상체계에 대한 공유.

- 선우, 성우, 장희 >> 프로젝트 관리 도서 Part2의 공부와 스터디!