**멘토 미팅 보고서**

|  |  |  |
| --- | --- | --- |
| 참여 인원  명단 | 김윤환 멘토 | 이예준 |
| 작성자 | 이예준 | |
| 미팅 주제 | Improvement of context-based meaning selection using Bi-Sent2Vec | |
| 미팅 일시 | 2021. 03. 31 | |
| 면담 내용 | 1. Preprocessing for training with Bi-Sent2Vec  - Bi-Sent2Vec으로 한국어-영어 문장 동시 Embedding을 시키기 위해 각 문장들을 형태소 단위로 분해하고 \_ko, \_en tagging하는 전처리 과정이 필요하다. 따라서 아래 사진과 같은 형식으로 총 pairs는 1,602,418개가 준비되었다. 형태소 분석을 위한 Khaiii 빌드와 각각의 corpus가 존재하는 엑셀파일을 읽어들이는 속도 등 꽤 오랜 시간이 걸렸다.  A picture containing text  Description automatically generated | |
| 증빙 | 매주 수요일마다 멘토와 오프라인으로 미팅하는 것으로 계획되어 있다.  Graphical user interface, text, application, chat or text message  Description automatically generated  3월 31일 오프라인 미팅 증빙 사진  A picture containing text, person, indoor, computer  Description automatically generatedA picture containing text, person, indoor, computer  Description automatically generated | |