|  |  |
| --- | --- |
| 교육 제목 | 자연어 처리 \_ 5 |
| 교육 일시 | 22.01.20 |
| 교육 장소 | 비대면 |
| **교육 내용** | |
| 오전 | 1.IMDB리뷰 Tensorflow의 SubwordTextEncoder로 토큰화  2.네이버 영화리뷰로 SubwordTokenizer |
| 오후 | 1.Bag Of Words(BOW)  2.Tensorflow의 Keras Tokenizer를 활요한 BOW  3.Scikit\_learn의 CountVectorizer를 활용한 BOW  4.불용어를 제거한 BOW만들기  5.사용자가 직업 정의한 불용어 사용하기  6.CountCectorizer에서 제공하는 자체 불용어 사용하기  7.NLTK에서 지원하는 불용어 사용하기 |