연구과제명 : 화자분리와 대본을 이용한 뉴스 음성 요약

|  |  |  |  |
| --- | --- | --- | --- |
| 화자의 중요도를 추출하기 위하여 문장 별로 중요도를 구하는 방법에 대하여 연구를 진행하였다. 지난 조사에서 text-rank에 word2vec를 적용하여 기존 text-rank에 비해서 성능 향상이 있었던 연구가 있었다. 이를 lex-rank에 적용하기 위해 문장을 벡터로 바꾸는 문장 임베딩에 대해서 조사를 하였다. 지난 조사에서 사용하던 word2vec의 경우 단어의 동시 출현을 기반으로 벡터를 생성하여 단어의 순서에 따른 문장의 맥락을 잘 반영하지 못하는 임베딩 결과가 나온다는 사실을 확인하였다. 그래서 이를 해결한 transformer기반의 임베딩 기법에 대한 조사를 실시하였다. transformer기반의 단어 임베딩 방법중 BERT기반의 모델의 성능을 비교해 보았다.  이 임베딩을 사용하여 코사인 유사도를 기반으로 rex-rank를 적용할 것이므로 관련 점수가 높은 모델을 선정하기로 하였다. 이와 관련된 지표인 cosine pearson과 cosine spearman중 코사인 유사도 기반 순위 상관 관계를 측정하는 spearman을 더 잘 계산하는 모델이 더 적합할 것이라고 판단하여 kosimCSE-RoBERTa-multitask모델이 적합할 것이라고 판단하였다. 이제 이 임베딩과 rex-rank를 사용하여 문장 별 중요도를 구하고, 화자 별 발화 길이와 문장 별 중요도를 종합적으로 고려하여 중심 화자를 정할 예정이다.  화자 분리의 경우 현재 countnet을 이용하여 화자 수를 추정하여 해당 구간에 몇 명의 화자가 있는지를 판단한 뒤, 그에 맞춰서 화자 분리를 진행하려고 하였다. 다만 해당 모델로 여러가지 실험을 하였는데 해당 모델이 general한 상황에서 정확한 결과가 나오지 않는다는 사실을 파악하였다. 일단 이 부분은 넘어가고 음성 구간 내에 몇 명의 화자가 있는지 직접 입력하는 방법으로 넘어가기로 한 뒤 이후 부분을 진행하였다. 분리된 음성을 화자 인식 모델을 통해 동일한 음성끼리 segment들을 연결하여 결과적으로 화자 별 음성을 분리하는 방법으로 시스템을 구현할 예정이다. 전체적인 흐름도를 대력적으로 나타내면 다음과 같다. | | | |
| **작성자** | **일자 2023-04-27** | **확인자** | **일자 2023-04-27** |
| **서명 신원철** | **서명 김유성** |