연구과제명: 화자분리와 대본을 이용한 뉴스 음성 요약

|  |  |  |  |
| --- | --- | --- | --- |
| 화자 별 요약을 생성하기 위해 few-shot learning을 진행하였다. 대본에서 화자 별로 어떤 내용의 발화를 하는지 따로 요약을 해 놓은 데이터 셋이 없기에 직접 화자 별 요약을 생성하고, 해당 데이터를 gpt-3의 few- shot learning의 입력으로 하고 다른 test data로 화자 별 요약의 결과를 확인 및 평가하였다. 데이터 셋의 경우 [전체 내용/화자 별 요약]의 내용이 포함되어 있고, 데이터의 일부는 아래와 같다.    위쪽 부분은 ‘전체 내용’ 부분이고, 아래쪽 부분은 ‘화자 별 요약’ 부분이다. 위와 같은 데이터를 ‘방송 콘텐츠 대본 요약 데이터’의 분류 기준에 따라 드라마, 시사, 역사, 문화 등의 분류에 따라 데이터를 만들고 해당 분유 별로 시스템의 성능을 측정하였다. 시스템의 성능은 rouge와 meteor를 기준으로 측정을 하였다. 하지만 화자 별 요약의 경우 기존 같은 작업을 하는 시스템의 부재로 인해 baseline과의 성능 비교는 할 수 없었다.    ‘문화’로 분류된 대본에 대한 시스템의 평가 결과는 다음과 같았다. | | | |
| **작성자** | **일자 2023-05-20** | **확인자** | **일자 2023-05-20** |
| **서명 신원철** | **서명 김유성** |