

컴퓨터공학부 캡스톤디자인 중간평가2 답변서

팀명: 1 조 안진마

조원: 신상훈, 김연수, 송성재, 박형준, 허진선, 윤정연

| 심사의견 or 질문 |
|---------------------------------------------------------------------------------------------------------------------------------------------------|
| 아직 연동이 안되어 있다고 했는데 왜 그런건가요? |
| 답변 |
| 당시 기준으로 프론트엔드와 백엔드 모두 대부분의 기능이 구현 되어있는 상태였으나 라벨링 관련 이슈를 해결하기 위해 모든 인력을 라벨링에 투입하였고 이 이슈를 마무리를 하지 못한 상태였습니다. 현재는 모두 연동된 상태이며 올바르게 작동하고 있습니다. |

| 심사의견 or 질문 |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 영상 정보를 이용하여 하이라이트 지점을 추출하는 것도 중요할것이라 생각되는데 이 부분은 왜 구현하지 않기로 했나요? |
| 답변 |
| 기존에 존재하는 데이터셋이 있어 이를 이용하여 영상 정보를 이용하여 하이라이트 지점을 추출하는 기능도 구현할 수 있다면 좋겠지만 이 기능을 구현하기에 적합한 데이터셋을 찾을 수 없었습니다. 직접 데이터셋을 구축하는 것에 대해 팀원들과 논의를 해본 결과 1인 크리에이터들의 콘텐츠(메이크업, 게임, 먹방 등)는 그 종류가 매우 다양하기 때문에 이 콘텐츠들을 종류별로 분류하고 각 카테고리별로 영상데이터의 라벨링을 진행하기에는 너무 많은 시간이 소요되어 진행에 어려움이 있다고 판단했습니다. 동영상의 하이라이트 예측에 대한 논문들을 찾아본 결과 채팅과 소리만으로 한 경우도 많았으며, 소리와 채팅은 영상과 비교했을때 폭넓은 콘텐츠들에 적용할 수 있기 때문에 영상 정보를 제외하고 소리와 채팅만을 이용하여 하이라이트 지점을 추출하기로 결정했습니다. 논문 링크 - 딥러닝을 이용한 양방향 소통 영상의 하이라이트 검출(김 은율) http://snut.dcollection.net/public_resource/pdf/200000220857_20200528181657.pdf |

| 심사의견 or 질문 |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p>채팅을 긍부정으로 분류할 때는 직접 데이터 라벨링을 진행하고 영상은 데이터셋이 존재하지 않다는 이유로 구현하지 않는 것은 일관성이 없습니다.</p> <p>데이터셋이 존재하지 않다고 하였지만 데이터셋도 있고 모델링도 존재합니다. 구글에서 제공하는 데이터셋 검색 사이트 (https://datasetsearch.research.google.com/)에서 필요한 영상 데이터셋이 있는지 찾아보세요.</p> |
| 답변 |
| <p>채팅에 관련하여 직접 데이터 라벨링을 한 이유는 채팅에 사용되는 문체 때문입니다. 자음만 사용한다거나 모음만 사용하는 등 일반적인 문체와 상이한 경우가 많기 때문에 네이버 영화 리뷰 데이터셋과 같은 기존에 존재하는 데이터셋을 이용하기보다는 직접 라벨링을 하는 것이 더 좋은 성능을 나타낼 것이라고 판단했습니다.</p> <p>또한, 알려주신 사이트에서 심도있게 검색해 보았으나 저희 프로젝트에 필요한 데이터셋은 찾지 못했습니다. 대부분 플랫폼의 유형 차트나 연령대별 차트 등의 분석 결과를 나타내고 있었으며 1인 크리에이터들의 영상에서 하이라이트 추천지점을 찾기에 관련된 데이터는 찾을 수 없었습니다.</p> <p>아래 링크를 통해 영상 데이터셋을 찾기 위해 사용한 검색어와 그 검색결과를 보실 수 있습니다. 검색 결과 저희 프로젝트에 필요한 데이터셋은 찾지 못했습니다. 데이터 조사.docx</p> |

| 심사의견 or 질문 |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p>감정을 긍정과 부정 2가지로만 분류하면 의미가 있을까요?</p> |
| 답변 |
| <p>컨텐츠 판단에 긍정과 부정이 가장 큰 척도가 되기 때문에 중요도를 높게 잡아 먼저 진행하게 됐습니다. 이후 7가지 감정 분석을 추가적으로 도입해 해당 컨텐츠를 시청자들이 긍정적으로 보는지, 부정적으로 보는지에 더불어 어떤 감정들로 인해 긍정 또는 부정적인 반응이 나왔는지 분석할 수 있습니다.</p> |

| 심사의견 or 질문 |
|-----------------------------------------------------------------------------------------------|
| <p>채팅을 긍정과 부정으로 분류하는 것은 가능하지만 이를 6가지 감정 다원화하여 분류하기 위해서는 다양한 라벨링 데이터가 필요합니다. 어떻게 분류할 것인가요?</p> |

| 답변 |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p>캐글에서 6가지 감정 분류에 대한 영문 텍스트 데이터셋 41만 라인을 찾았으나 문체가 채팅과 상당히 다른 양상을 가지는 것 뿐만 아니라 영문을 한글로 번역하여 사용할 경우 지나치게 번역체이며 직독되어 문장의 의미가 퇴색되는 문제가 있었습니다.</p> <p>따라서 다시보기 영상의 채팅로그에서 다운받은 채팅들에 대해 각각 라벨링을 진행하여 자체 데이터셋을 구축한 뒤 감정분류모델에 사용하였습니다.</p> <p>이 데이터셋은 기존 긍정/부정 분류에 사용한 데이터셋과는 다른 데이터셋으로 joy, sadness, love, fear, surprise, anger, neutral 총 7가지 부류로라벨링 되어 있습니다.</p> |

| 심사의견 or 질문 |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p>구축한 솔루션에 신뢰성이 의심됩니다. 자체적으로 제작한 데이터셋의 객관성과 신뢰성이 떨어지는 것 같습니다.</p> |
| 답변 |
| <p>객관성과 신뢰성을 높이기 위해 다음과 같은 작업을 진행했습니다.</p> <p>소리와 채팅으로 하이라이트 지점을 추출 할 수 있다는 것을 증명하기 위해 편집된 하이라이트 영상이 존재하는 60개 스트리밍 동영상에서 하이라이트 지점을 추출한 후 스트리밍 동영상에서 파생된 하이라이트 동영상에 추출된 지점이 존재하는지 확인하는 작업을 진행하였습니다. 그 결과 추출된 하이라이트 지점 중 약 75.8%가 실제로 하이라이트 영상에 존재함을 확인하였습니다.</p> <p>해당 내용을 정리한 문서는 깃에 올려놨습니다. (/doc/2_중간 자문 평가/휴리스틱 타당성 검증 보고서-요바(요기 바바)-1조.pdf)</p> <p>또한, 긍정 부정 데이터셋을 구축할 때에는 신뢰성을 위해 스탠포드 대학에서 사용하는 라벨링 방식을 이용하여 약 20000개의 데이터를 5명에서 동시에 라벨링 진행하였습니다. 이때 1(긍정) 또는 0(부정)으로 라벨링하여 5명의 평균값 0.5 이상이면 긍정, 미만이면 부정으로 최종 판단하여 라벨링을 진행하였습니다.</p> |

| 심사의견 or 질문 |
|-----------------------------------------------------------------------|
| <p>채팅의 긍부정 분류의 정확도를 측정하였는데 학생들이 라벨링한게 정답이고 그거에 얼마나 맞는건지를 계산한 건가요?</p> |

| 답변 |
|--------------------------------------------------------------------------------------------------------------------------------------------|
| 직접 구축한 2만개의 채팅 라벨링 데이터를 15000개는 train-set, 5000개는 test-set으로 사용하여 정확도를 측정하였습니다. 실험 결과 Okt보다 SPM 토큰라이저를 사용하는 것이 채팅 공부정 분류에 더 높은 정확성을 보였습니다. |

| 심사의견 or 질문 |
|-----------------------------------------------------------------|
| 채팅 데이터는 어디서 가져왔나요? |
| 답변 |
| 아프리카TV, 유튜브, 트위치에 스트리밍한 동영상의 채팅 기록을 다운로드 하는 모듈을 직접 구현하여 수집했습니다. |

| 심사의견 or 질문 |
|----------------------------------------------------------------------------------------------------|
| 어떤 기준으로 몇 개의 동영상을 선별하여 채팅 데이터셋을 만들었나요? |
| 답변 |
| 아프리카TV, 유튜브, 트위치에 스트리밍한 동영상 중에서 기록된 채팅 수가 많은 동영상들을 선택하여 2만개의 채팅 데이터를 수집한 후 이를 라벨링하여 데이터셋을 구축하였습니다. |

| 심사의견 or 질문 |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 채팅에서 Kkma 형태소 분석기를 이용하여 명사를 추출한 결과를 보면 '기훈'과 '기훈이'와 같이 조사의 유무의 차이만 있는 단어가 각각 추출되는데 이런 것들은 '기훈' 하나로 인식하여 추출되도록 해야하는 것이 아닌가요? |
| 답변 |
| <p>발표의 전달력이 부족했었던 것 같습니다.</p> <p>해당 표는 Komoran, Kkma, Hananum, Okt 형태소 분석기를 이용하여 명사를 추출한 결과를 나타낸 표로 명사를 추출하는 데 어떤 형태소 분석기를 사용하는 것이 더 좋을지를 판단하기 위해 진행한 과정을 보여드리기 위해 작성된 표입니다.</p> <p>형태소 분석기마다 명사를 추출한 결과가 서로 상이했고 이를 비교해본 결과 초성(ㅋㅋㅋ, ㅎㅎ 등)을 명사로 인식하지 않고 다른 형태소 분석기보다 조사를 분리하여 명사를 추출하는 능력이 뛰어났던 Okt 형태소 분석기가 명사를</p> |

추출하기에 가장 적합하다고 판단했다는 것을 알려드리기 위해 해당 표를 보여드렸습니다.

‘기훈이’, ‘기훈’ 이렇게 두가지 명사를 추출한 형태소 분석기는 Kkma이며 위에 설명한 과정을 통해 Okt 형태소 분석기를 사용하는 것이 가장 적합하다고 판단하여 프로젝트에서는 Okt 형태소 분석기를 사용하였습니다.