 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	인공지능을 이용한 하이라이트 추천 시스템	
	팀 명	안진마	
	Confidential Restricted	Version 2.3	2020-APR-23

# 캡스톤 디자인 I

## 종합설계 프로젝트


프로젝트 명	요바(요기 바바)
팀 명	안진마
문서 제목	인공지능을 이용한 하이라이트 추천 시스템 계획서

Version	2.3
Date	2020-APR-23

팀원	신상훈 (조장)
	김연수
	송성재
	박형준
	허진선
	윤정연

### CONFIDENTIALITY/SECURITY WARNING


이 문서에 포함되어 있는 정보는 국민대학교 전자정보통신대학 컴퓨터공학부 및 컴퓨터공학부 개설 교과목 캡스톤 디자인 수강 학생 중 프로젝트 요바를 수행하는 팀 안진마의 팀원들의 자산입니다. 국민대학교 컴퓨터공학부 및 팀 안진마의 팀원들의 서면 허락없이 사용되거나, 재가공 될 수 없습니다.

 <b>국민대학교 컴퓨터공학부 캡스톤 디자인 I</b>	<b>계획서</b>		
	<b>프로젝트 명</b>	인공지능을 이용한 하이라이트 추천 시스템	
	<b>팀 명</b>	안진마	
	Confidential Restricted	Version 2.3	2020-APR-23

## 문서 정보 / 수정 내역


<b>Filename</b>	계획서-요바(요기바바).pdf
<b>원안작성자</b>	신상훈
<b>수정작성자</b>	신상훈, 김연수, 송성재, 박형준, 허진선, 윤정연

수정날짜	대표수정자	Revision	추가/수정 항목	내 용
2020-03-09	신상훈	1.0	최초 작성	목차별 초안 및 개요, 추진 배경
2020-03-11	박형준	1.1	프로젝트 비용	프로젝트 비용
2020-03-14	김연수	1.2	목표 작성	개발 세부 목표
2020-03-17	송성재	1.3	배경기술 작성	프로젝트 배경기술 및 제한요소
2020-03-19	허진선	1.4	기대효과 작성	기대효과 및 활용방안
2020-03-20	윤정연	1.5	요구사항 작성	기능적, 비기능적 요구사항과 품질 속성
2020-03-23	전원	1.6	피드백 및 수정	피드백 및 수정
2020-03-26	신상훈	1.7	최종 검토	최종 검토
2020-04-18	김연수	2.0	연구/개발내용 수정	연구/개발내용 수정
2020-04-18	허진선	2.1	시스템 구조, Use case	시스템 구조, Use case 작성
2020-04-20	허진선	2.2	Actor-goal list 작성	Actor-goal list 작성
2020-04-23	김연수	2.3	시스템 구성도	시스템 구성도 작성
2020-04-23	신상훈	2.3	최종 검토	최종 검토

 <b>국민대학교 컴퓨터공학부 캡스톤 디자인 I</b>	<b>계획서</b>		
	<b>프로젝트 명</b>	인공지능을 이용한 하이라이트 추천 시스템	
	<b>팀 명</b>	안진마	
	Confidential Restricted	Version 2.3	2020-APR-23

## 목 차

<b>개요</b>	<b>4</b>
프로젝트 개요	4
추진 배경 및 필요성	5
<b>개발 목표 및 내용</b>	<b>7</b>
목표	7
연구/개발 내용	8
개발 결과	12
시스템 기능 요구사항	12
시스템 비기능(품질) 요구사항	16
시스템 구조	17
결과물 목록 및 상세 사양	20
기대효과 및 활용방안	20
<b>배경 기술</b>	<b>21</b>
기술적 요구사항	21
현실적 제한 요소 및 그 해결 방안	21
하드웨어	21
소프트웨어	21
기타	22
<b>프로젝트 팀 구성 및 역할 분담</b>	<b>23</b>
<b>프로젝트 비용</b>	<b>24</b>
<b>개발 일정 및 자원 관리</b>	<b>25</b>
개발 일정	25
일정별 주요 산출물	26
인력자원 투입계획	28
비 인적자원 투입계획	28
<b>참고 문헌</b>	<b>29</b>

 <b>국민대학교 컴퓨터공학부 캡스톤 디자인 I</b>	<b>계획서</b>		
	<b>프로젝트 명</b>	인공지능을 이용한 하이라이트 추천 시스템	
	<b>팀 명</b>	안진마	
	Confidential Restricted	Version 2.3	2020-APR-23

# 1 개요

## 1.1 프로젝트 개요


교육부에서 발표한 ‘2019년 진로교육 현황조사’에 의하면 크리에이터(Creator)가 희망직업 순위 3위인 시대다.[그림 1-1] 2018년부터 통계청에선 이 직업을 ‘미디어 콘텐츠 창작자’라는 직업군으로 분류하기 시작했고, 중앙일보 보도자료에 따르면 유튜브(Youtube)에 1분마다 업로드 되는 영상이 400시간에 달한다고 한다. 이런 시대인 만큼 크리에이터에 도전하는 사람이 많아지고 있는 추세이다.



[그림 1-1] 초등학생 희망직업 순위, 중앙일보

많은 1인 크리에이터들은 실시간 인터넷 방송을 진행 후 장시간의 방송 장면 중에서 하이라이트를 선별한다. 여기에 다양한 효과를 첨가하여 편집한 영상을 유튜브나 기타 플랫폼에 게시하면, 이를 통해 시청자를 늘리고 수익을 창출할 수 있다.

유명 크리에이터의 경우 기획자·편집자들과 공동으로 작업을 하지만, 신규 크리에이터들은 대부분 혼자 힘으로 기획부터 편집까지 해야하는 경우가 많다. 여러 플랫폼의 다시보기 카테고리를 보면 대부분 오랜 시간 방송을 진행하는 것을 알 수 있다. 이런 긴 영상을 재미있는 장면만을 골라 보기 좋은 짧은 영상으로 편집하기 위해선 해당 하이라이트 부분을 찾기 위해 영상을 여러번 돌려봐야할 가능성이 크다. 시간이 한정된 상황에서 편집에 소요되는 시간이 길어질수록 해당 크리에이터는 콘텐츠 제작뿐만 아니라 전체적인 결과물의 질이 나빠질 수 있다. 이러한 사항들은 신규 크리에이터들 뿐만아니라 편집자들에게도 해당되는데, 질 좋은 편집을

 <b>국민대학교 컴퓨터공학부 캡스톤 디자인 I</b>	<b>계획서</b>		
	<b>프로젝트 명</b>	인공지능을 이용한 하이라이트 추천 시스템	
	<b>팀 명</b>	안진마	
	Confidential Restricted	Version 2.3	2020-APR-23

하기 위해서는 하이라이트 지점 탐색 시간 감축이 필요하다. 이는 질 높은 영상 제작에 큰 기여를 할 것이다.


영상을 편집할 때 편집자는 하이라이트 지점을 찾기 위해 많은 것들을 고려해야 한다. 이런 수고를 덜어주기 위해 영상의 특징들을 파악하여 하이라이트 수준이라고 판단되는 지점을 크리에이터나 편집자에게 하이라이트 지점으로 추천해주고 나아가 어떤 행동이 있었을때 시청자의 반응이 긍정적이었는지 부정적이었는지도 알려주는 서비스를 제공하는 것이다.

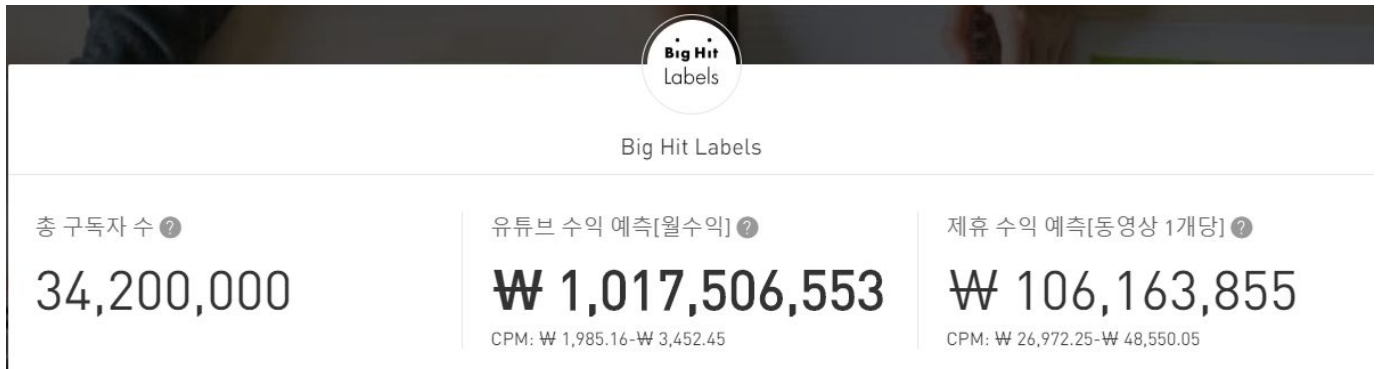
## 1.2 추진 배경 및 필요성

와이즈앱에 의하면 현재 전연령에 걸쳐 앱 사용량을 조사해본 결과 유튜브가 1위이고 4위는 페이스북이다.[그림1-2] 두 어플 모두 주로 10분 안팎으로 편집된 영상이 업로드되는 플랫폼으로 사람들은 이동하면서, 혹은 여가시간에 이를 시청한다. 그리고 크리에이터들은 이 플랫폼들을 이용하여 광고 및 기타 사항으로 수익을 재창출 한다. 유튜브 통계 사이트인 Nox Influencer에 따르면 가장 많은 구독자를 보유하고 있는 Big hit Label 채널의 경우 영상 한 개당 수익이 억 단위를 넘어간 것으로 추정된다.[그림 1-3] 이처럼 본 프로젝트가 타겟으로한 시장의 크기가 매우 큼을 알 수 있다.



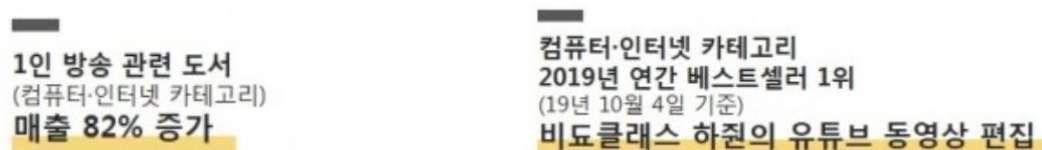
[그림 1-2] 한국인이 오래 사용하는 앱, 와이즈앱

 <b>국민대학교 컴퓨터공학부 캡스톤 디자인 I</b>	<b>계획서</b>		
	<b>프로젝트 명</b>	인공지능을 이용한 하이라이트 추천 시스템	
	<b>팀 명</b>	안진마	
	Confidential Restricted	Version 2.3	2020-APR-23



[그림 1-3] Big Hit Labels 채널의 유튜브 수익 예측, Nox Influencer

10분 내외의 영상이 주로 업로드되는 플랫폼의 성장에 따라, 크리에이터들에게 편집 작업은 큰 고민거리가 되어 다가온다. 온라인 서점인 인터파크의 통계에 따르면 편집관련 서적이 연간 베스트 셀러 1위를 달성하고 판매량이 82% 증가했다고 한다.[그림 1-4][그림 1-5] 이와 같이, 크리에이터들에게 편집은 매우 중요하며, 1인 방송 시장이 활발해 지고 있는 만큼 이들을 타겟으로한 서비스가 필요하다고 할 수 있다.




[그림 1-4] 1인 방송 관련 도서 매출, 인터파크 [그림 1-5] 2019년 연간 베스트셀러 1위, 인터파크

실력 능숙하게 편집을 한다해도 하이라이트 지점을 찾기 위해 투자하는 시간을 줄이지 못한다면 업로드 시간이 불규칙하게 되어 시청자와의 갈등이 생길 가능성이 크다. 하이라이트 지점을 추천해주고 시청자들의 목소리인 채팅을 분석해주는 것으로 영상 편집 시간을 줄이고 본인 영상에 대한 피드백도 받을 수 있다면 크리에이터와 편집자에게 무엇보다도 값진 서비스가 될 것이다.

또한 마케팅 전략으로 유튜브나 기타 플랫폼 광고가 빠질 수 없는 현 시점에서 광고주들은 시청자들의 반응이나 광고를 붙이고 싶은 위치를 모색하는데 본 서비스를 이용한다면 보다 좋은 광고 효과를 얻을 수 있을 것이다. 말 그대로 확장성이 큰 사업이 될 것이다.

기존의 인공지능을 사용한 하이라이트 영상을 보면 대부분이 스포츠나 아이돌 영상에 포커스가 맞춰져 있다. 한 선수나 인물만을 찾아서 편집 영상을 만들어 주는 것과 관중의 함성 및 골을 넣는 장면을 학습시켜 중요 장면을 보여주는 것에 그치고 있다. 하지만 인터넷 방송은 스포츠와 달리 하이라이트를 판별하기 위한 관중도 골 장면도 없다. 따라서 1인 크리에이터를 타겟으로 설정한 본 프로젝트는 기존에 출시된 서비스도 없을 뿐더러 확장성이 큼을 알 수 있다.

 <b>국민대학교 컴퓨터공학부 캡스톤 디자인 I</b>	<b>계획서</b>		
	<b>프로젝트 명</b>	인공지능을 이용한 하이라이트 추천 시스템	
	<b>팀 명</b>	안진마	
	Confidential Restricted	Version 2.3	2020-APR-23

## 2 개발 목표 및 내용

### 2.1 목표

본 프로젝트는 인터넷 방송 크리에이터가 장시간의 방송을 마친 다시보기 영상을 분석하여 하이라이트 추천지점을 제공한다.

시장조사 결과 현재 인터넷 실시간 방송 국내 시청자 선호 플랫폼 중 상위 3개에 랭크된 플랫폼은 유튜브, 아프리카TV, 트위치인 것으로 확인된다.[그림 2-1] 이에 따라 본 프로젝트는 해당 3사를 타겟으로 서비스를 기획했다.




[그림 2-1] 국내 시청자 선호 인터넷 방송 플랫폼, 나스미디어

필수 개발 목표로는 다음이 있다.

1. 동영상의 소리를 추출하여 초당 데시벨을 측정하고 기록할 수 있다.
2. 채팅로그를 이용하여 시청자의 댓글 참여가 많았던 포인트들을 탐색하고 기록할 수 있다.
3. 채팅로그를 이용하여 채팅에 참여한 시청자 수를 측정하고 기록할 수 있다.
4. 채팅로그를 이용하여 시청자들의 긍부정 반응을 체크할 수 있다.
5. 크리에이터 또는 편집자가 동영상 링크를 제출하고 분석 결과를 리턴받을 웹 페이지를 작성한다.

부가적인 목표로 다음을 제공한다.

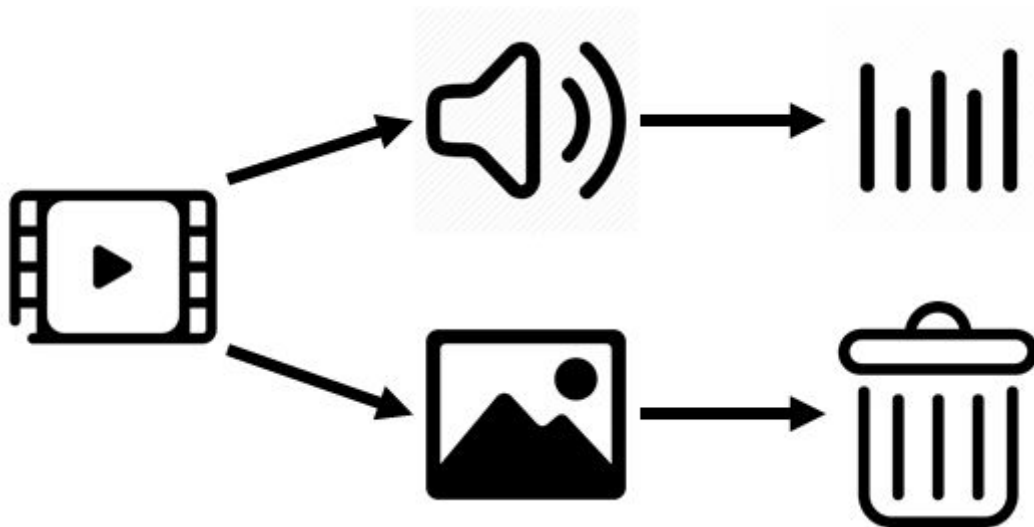
1. 시청자의 반응과 영상을 종합적으로 분석하여 시청자 반응성 그래프를 제공한다. 이는 크리에이터의 캐릭터를 찾는 데 도움을 준다.
2. 하이라이트 추천지점 뿐 아니라 시청자의 참여도(채팅, 도네이션 등)를 가시화 해서 나타낸다. 이는 편집자가 풀 영상을 시청하지 않고 선택적으로 시청할 수 있도록 해준다.
3. 크리에이터의 영상 분석 데이터를 저장할 데이터베이스를 생성한다. 이전에 분석했던 내역을 즉각적으로 불러올 수 있고 변화량을 관찰 할 수 있게 해준다.

 <b>국민대학교 컴퓨터공학부 캡스톤 디자인 I</b>	<b>계획서</b>		
	<b>프로젝트 명</b>	인공지능을 이용한 하이라이트 추천 시스템	
	<b>팀 명</b>	안진마	
	Confidential Restricted	Version 2.3	2020-APR-23

## 2.2 연구/개발 내용


### 1. 음성파일 분류 모듈

1인방송의 특성상 자신을 촬영해주는 사람이 따로 없기 때문에 카메라가 고정된 위치에 있을 수 밖에 없다. 그렇기 때문에 크리에이터는 한정된 공간밖에 사용하지 못하기 때문에 신체적인 움직임 보다는 목소리를 이용해 자신의 감정을 나타내고 시청자들과 공감하며 활동성을 나타내기 위해 집중한다. 따라서 크리에이터의 활동성을 분석하는데 오디오 파일을 사용하기로 하였으며 동영상 파일은 크기가 매우 크고 불필요 하기 때문에 걸러내는 작업이 필요하다. 추출된 오디오 파일은 스펙트럼을 분석하여 하이라이트 추천 알고리즘에 사용할 수 있도록 수치화한다.



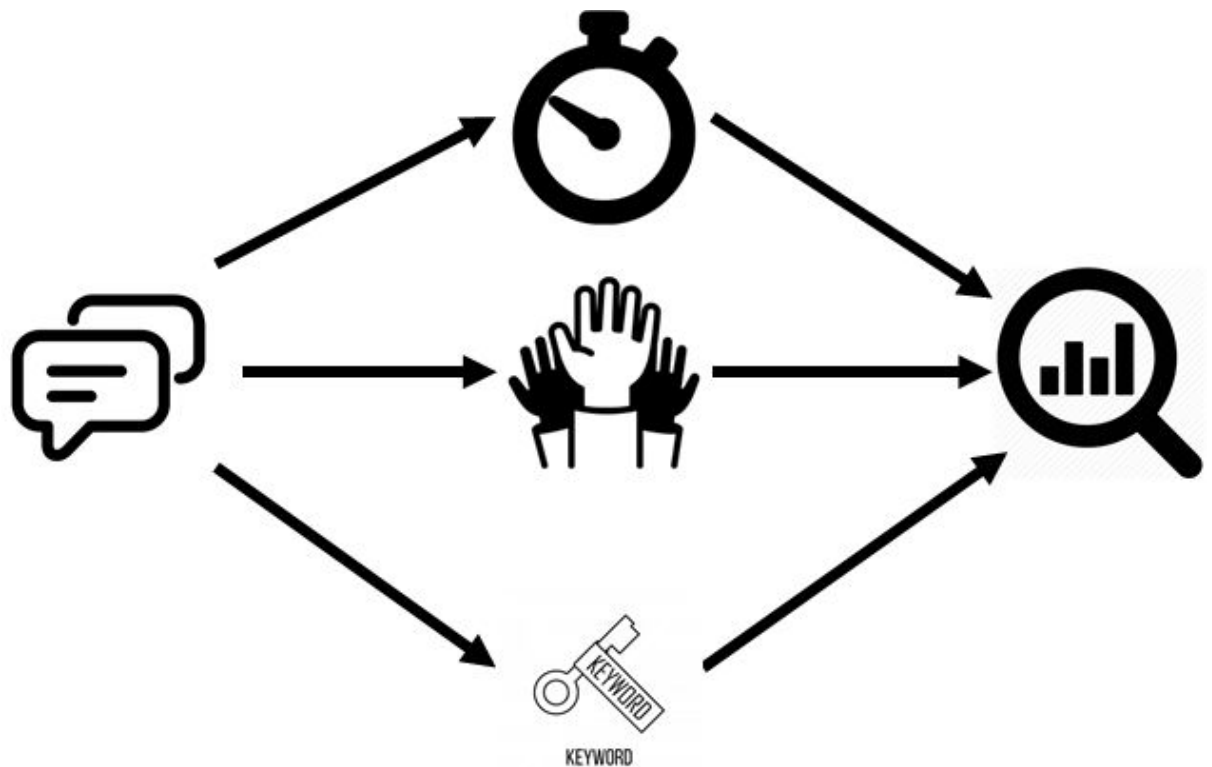
[그림 2-2] 음성파일 분류 모듈




 <b>국민대학교 컴퓨터공학부 캡스톤 디자인 I</b>	<b>계획서</b>		
	<b>프로젝트 명</b>	인공지능을 이용한 하이라이트 추천 시스템	
	<b>팀 명</b>	안진마	
	Confidential Restricted	Version 2.3	2020-APR-23

## 2. 채팅 로그 분석 모듈

크리에이터는 실시간 스트리밍을 통해 시청자와 즉각 소통하며 영상을 제작할 수 있다. 시청자의 채팅을 읽고 수행하는 콘텐츠도 있을 정도로 크리에이터들은 시청자들의 참여를 꾸준히 유도한다. 즉 시청자는 이제 제작된 영상을 소비하기만 하는것이 아니라 영상의 제작에도 참여하는 소비자이자 제작참여자 가 된 것이다. 때문에 크리에이터와 편집자에게 시청자의 반응을 체크하는 것은 선택이 아닌 필수가 됐다. 그래서 채팅로그를 추출하여 채팅 시간을 이용하여 시간대별 채팅 수를 분석하고, 채팅 아이디를 이용하여 채팅에 참여한 시청자 수를 분석한다. 또한 형태소 분석을 통해 자주 출현하는 명사를 키워드로 추출 하고 이 분석 내용들을 수치화한다.

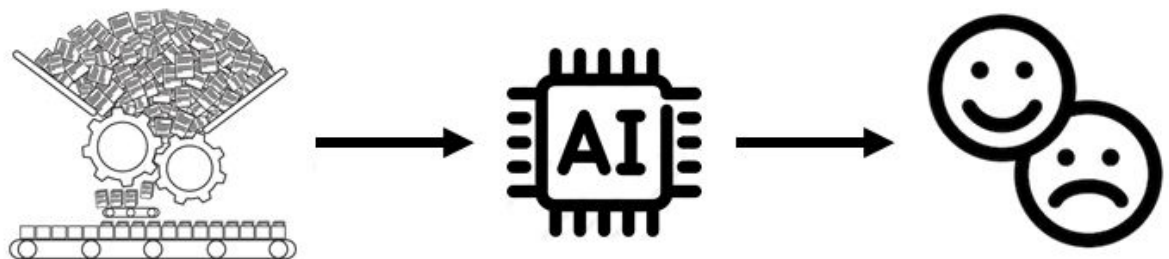


[그림 2-3] 채팅 로그 분석 모듈

 <b>국민대학교 컴퓨터공학부 캡스톤 디자인 I</b>	<b>계획서</b>		
	<b>프로젝트 명</b>	인공지능을 이용한 하이라이트 추천 시스템	
	<b>팀 명</b>	안진마	
	Confidential Restricted	Version 2.3	2020-APR-23

### 3. 긍부정 감정 분석 모듈

시청자들의 반응에는 절대적인 채팅량만이 중요한 것이 아니다. 만약 특정 구간에서 채팅량이 폭발적으로 증가했지만 대다수가 부정적인 댓글이라면 이 구간은 하이라이트 지점으로 선정해서는 안 될 것이다. 따라서 인공지능을 통해 채팅 내용을 감정 분석을 해 줄 필요가 있다. 이를 위해 인공지능 모델을 구축 하고, 150,000개의 네이버 영화 리뷰 데이터셋과 채팅 로그 15,000개를 라벨링 한 데이터를 이용해 모델을 학습 시켜 채팅을 감정 분석 한 뒤 수치화 한다. 긍부정 모델이 성공적으로 아웃풋을 뽑아 낸다면 차후 긍부정 뿐 아니라 더 다양한 감정을 분석 할 수 있도록 기능을 더할 예정이다.




[그림 2-4] 긍부정 감정 분석 모듈

### 4. 사용자가 서비스를 쉽게 이용할 수 있도록 웹 페이지를 구성한다.

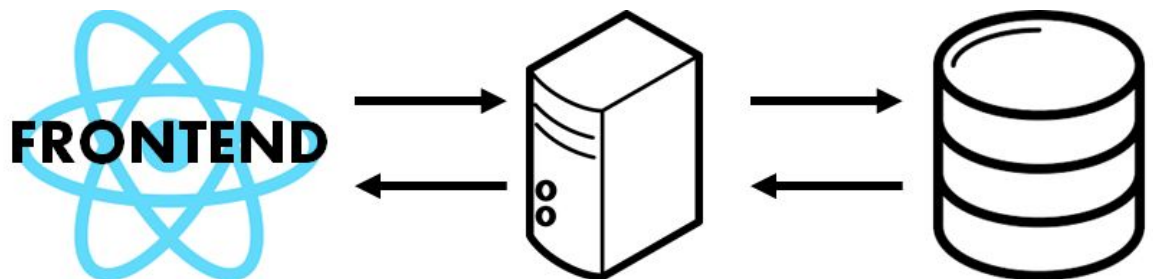
우리의 서비스는 편리함을 추구한다. 따라서 프로그램 설치 없이 간편하게 사용할 수 있도록 웹으로 제작한다. 또한 메뉴얼을 읽지 않고도 자연스럽게 사용할 수 있도록 UI / UX 디자인을 적용해 사용성을 높이려고 한다. 분석이 종료된 이후에는 수학적 지식이 많지 않더라도 쉽게 읽을 수 있는 그래프 종류를 이용하여 가시화하여 보여줄 것이다.




[그림 2-5] 웹 페이지

 <b>국민대학교 컴퓨터공학부 캡스톤 디자인 I</b>	<b>계획서</b>		
	<b>프로젝트 명</b>	인공지능을 이용한 하이라이트 추천 시스템	
	<b>팀 명</b>	안진마	
	Confidential Restricted	Version 2.3	2020-APR-23

5. 서비스를 제공할 수 있는 서버를 구성한다.
- 웹 페이지에서 정적인 화면을 보여줄 것이 아니기 때문에 서버는 필수적으로 구성하게 된다. 프레임워크는 플라스크를 사용할 것이며 프론트 서버와 연동을 위해 CORS를 적용할 것이다. 또한 웹 페이지가 구성이 되었지 않아도 서버를 개발 할 수 있도록 테스트 코드를 작성하며 API를 개발하려고 한다. 관계형 데이터베이스인 Postgresql을 사용하고 ORM은 SQLAlchemy를 사용할 것이며 DB 스키마는 alembic을 이용하여 버전관리를 할 것이다.



[그림 2-6] API 연결 구성


 <b>국민대학교 컴퓨터공학부 캡스톤 디자인 I</b>	<b>계획서</b>		
	<b>프로젝트 명</b>	인공지능을 이용한 하이라이트 추천 시스템	
	<b>팀 명</b>	안진마	
	Confidential Restricted	Version 2.3	2020-APR-23

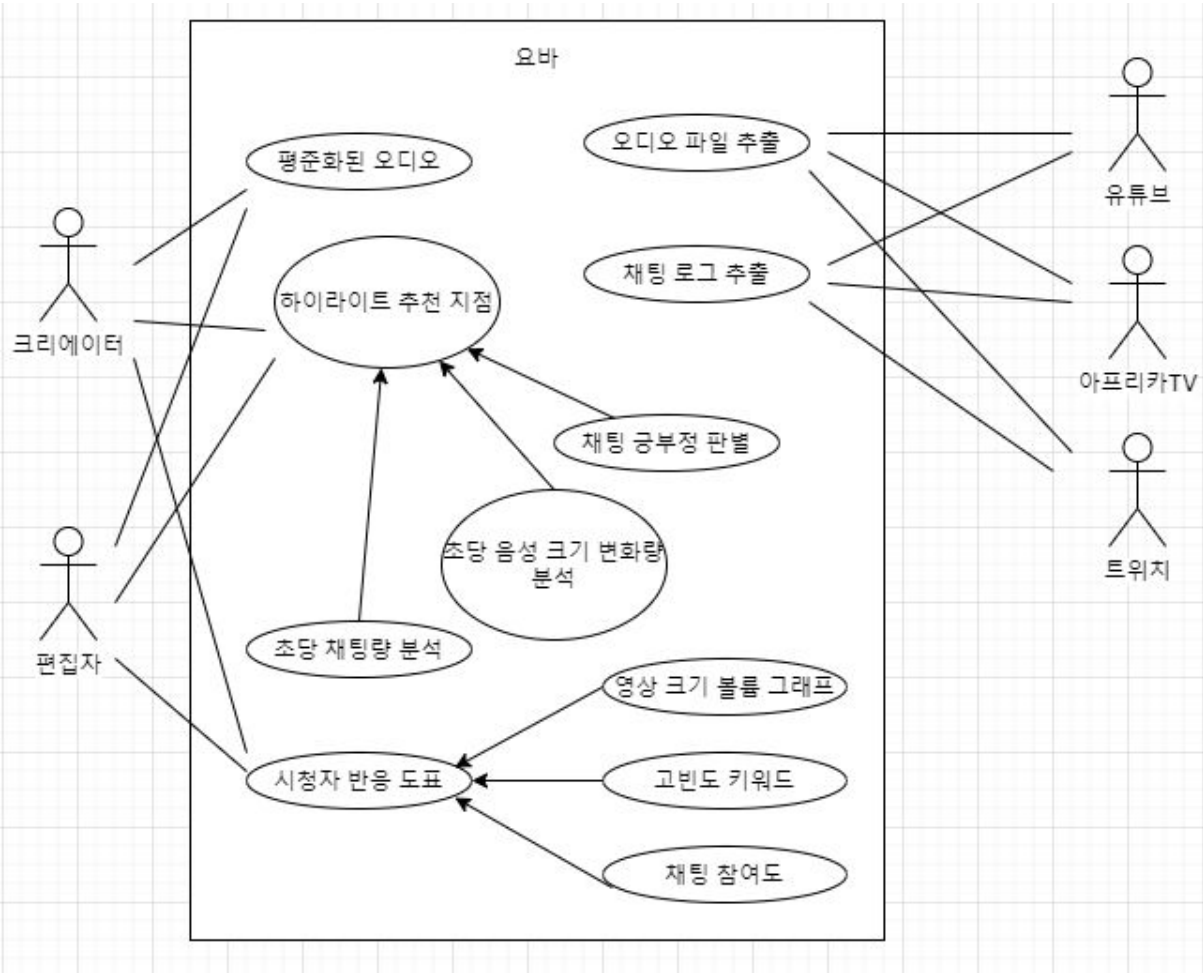
## 2.3 개발 결과

### 2.3.1 시스템 기능 요구사항


Functional Requirement	내용	중요도 (상/중/하)
FR 1	사용자가 원하는 플랫폼의 다시보기 URL을 입력할 수 있다.	상
FR 2	시스템은 입력받은 URL을 통해 오디오 파일 및 채팅 로그를 가져온다.	상
FR 5	추출 된 소리를 이용하여 초당 평균 음량을 계산한다.	상
FR 6	채팅 로그를 이용하여 초당 채팅 수를 계산한다.	상
FR 7	채팅 로그를 이용하여 시간대 별 시청자들의 채팅량을 통계낸다.	중
FR 8	채팅 로그를 이용하여 시청자들의 반응을 긍정/부정으로 분류한다.	하
FR 9	사용자에게 실행 결과를 보여주기 위한 UI를 제공한다.	상
FR 10	사용자가 웹 브라우저에서 프로그램을 이용할 수 있게 서버를 이용한다.	상
FR 11	사용자가 적절하지 않은 URL을 입력하면 에러 메시지를 출력한다.	상
FR 12	사용자가 원한다면 분석 도중 취소할 수 있다.	하

[표 2-1] 시스템 기능적 요구사항

 <div> <b>국민대학교</b>  <b>컴퓨터공학부</b>  <b>캡스톤 디자인 I</b> </div>	<b>계획서</b>		
	<b>프로젝트 명</b>	인공지능을 이용한 하이라이트 추천 시스템	
	<b>팀 명</b>	안진마	
	Confidential Restricted	Version 2.3	2020-APR-23



[그림 2-7] 유즈케이스 다이어그램


 <b>국민대학교 컴퓨터공학부 캡스톤 디자인 I</b>	<b>계획서</b>		
	<b>프로젝트 명</b>	인공지능을 이용한 하이라이트 추천 시스템	
	<b>팀 명</b>	안진마	
	Confidential Restricted	Version 2.3	2020-APR-23

Actor		Goal
Primary actor	크리에이터	1. 하이라이트 지점을 추천 받는다. 2. 시청자의 반응을 시각화 된 도표로 제공받는다. 3. 평준화 된 오디오 파일을 다운받는다.
	편집자	1. 하이라이트 지점을 추천 받는다. 2. 시청자의 반응을 시각화 된 도표로 제공받는다. 3. 평준화 된 오디오 파일을 다운받는다.
Offstage actor	시청자	1. 하이라이트 지점을 추천 받는다.
Supporting actor	Yoba	1. URL을 입력받아 Youtube, Twitch, AfreecaTV 중 적절한 데이터를(오디오, 채팅로그) 다운 받는다. 2. 하이라이트 지점을 추출한다. 3. 데이터를 분석하고 시각화 한다. 4. 오디오 볼륨 크기를 평준화한다.

[표 2-2] Actor-goal list

유스케이스명	URL 입력
트리거	사용자가 URL을 입력한다.
목표	입력된 URL을 이용해 동영상에 액세스 할 수 있다.
과정	1. 사용자가 분석하고 싶은 동영상의 URL을 입력한다. 2. 서비스는 입력받은 URL이 유효한지 정규식을 이용해 검사한다. a. 올바르지 않을 경우 에러메시지를 리턴한다. b. 올바른 경우 다음 과정을 진행한다. 3. 입력받은 URL을 이용해 동영상에 액세스한다.

[표 2-3] Usecase1

 <b>국민대학교 컴퓨터공학부 캡스톤 디자인 I</b>	<b>계획서</b>		
	<b>프로젝트 명</b>	인공지능을 이용한 하이라이트 추천 시스템	
	<b>팀 명</b>	안진마	
	Confidential Restricted	Version 2.3	2020-APR-23

유스케이스명	오디오파일 & 채팅로그 추출
트리거	액세스한 동영상 이 가용상태이다.
목표	오디오파일과 채팅로그를 추출하여 분석 가능한 상태로 만든다.
과정	<ol style="list-style-type: none"> <li>1. 액세스한 영상에서 moviepy를 이용하여 오디오파일을 추출한다.</li> <li>2. 액세스한 영상에서 beautiful soup 또는 자체적으로 제공하는 API를 이용하여 채팅로그를 추출한다.</li> </ol>


[표 2-4] Usecase2

유스케이스명	오디오파일 & 채팅로그 분석
트리거	오디오파일과 채팅로그추출에 성공한다.
목표	분석 결과를 도표화 할 수 있도록 수치화한다.
과정	<ol style="list-style-type: none"> <li>1. Moviepy를 이용하여 음성파일의 데시벨을 이용하여 분석한다.</li> <li>2. 알고리즘을 통하여 시간대별, 구간별 채팅량, 고빈도로 등장하는 키워드, 시청자의 채팅 참여도를 통계낸다.</li> <li>3. 인공지능을 이용하여 채팅을 감정분석하여 긍정성 수치를 낸다.</li> <li>4. 위 1 ~ 3 분석 결과를 통해 하이라이트 추천지점을 생성한다.</li> <li>5. 모든 과정이 완료되면 프론트엔드로 리턴한다.</li> </ol>

[표 2-5] Usecase3

유스케이스명	하이라이트 지점 추천 & 도표화
트리거	데이터 분석 종료.
목표	데이터 분석 결과를 이용하여 사용자에게 가시화해준다.
과정	<ol style="list-style-type: none"> <li>1. 분석결과를 라이브러리를 이용하여 도표화하여 사용자에게 제공한다.</li> <li>2. 리턴받은 하이라이트 추천지점을 사용자에게 제공한다.</li> </ol>

[표 2-6] Usecase4


 <b>국민대학교 컴퓨터공학부 캡스톤 디자인 I</b>	<b>계획서</b>		
	<b>프로젝트 명</b>	인공지능을 이용한 하이라이트 추천 시스템	
	<b>팀 명</b>	안진마	
	Confidential Restricted	Version 2.3	2020-APR-23

### 2.3.2 시스템 비기능(품질) 요구사항

Nonfunctional Requirement	내용	중요도 (상/중/하)	품질 속성 (ISO 25010)
NFR 1	프로그램 실행 시간을 5분 이내로 한다.	중	실행효율성-시간효율성
NFR 2	기능들을 가능한 최소의 단위(함수)로 구성한다.	중	유지보수성-모듈성
NFR 3	유튜브뿐 아니라 아프리카TV, 트위치의 다시보기 URL도 사용 가능하게 한다.	상	사용성-접근성
NFR 4	사용자가 불편함을 느끼지 않도록 사용자 중심의 UI/UX를 구성한다.	중	사용성-이해용이성
NFR 5	분석한 결과들을 데이터베이스에 저장하여 재분석 없이 바로 불러올 수 있게한다.	중	사용성-운영성
NFR 6	별도의 설치 없이 사용할 수 있도록 웹으로 프로그램을 구성한다.	상	이식성-치환성
NFR 7	사용자가 어려움 없이 인터페이스를 이용할 수 있도록 도움말을 삽입한다.	하	사용성-학습성
NFR 8	AWS를 이용하여 접속자 수와 관계없이 시스템이 유지되도록 한다.	중	신뢰성-가용성

[표 2-7] 시스템 비기능적 요구사항



 <b>국민대학교 컴퓨터공학부 캡스톤 디자인 I</b>	<b>계획서</b>		
	<b>프로젝트명</b>	인공지능을 이용한 하이라이트 추천 시스템	
	<b>팀명</b>	안진마	
	Confidential Restricted	Version 2.3	2020-APR-23


### 2.3.3 시스템 구조

#### 프론트엔드

- 1) 화면 구성 컴포넌트 모듈 ( React를 통한 View 모듈 제작 )
  - a) Navibar, Footer, Main 컴포넌트 등을 구성해 기본 웹페이지 구축
  - b) 백엔드에서 오디오와 채팅을 분석한 결과 API를 호출해 가시화 및 다운로드 컴포넌트 구현
    - i) 하이라이트 지점 지표화
    - ii) 채팅 결과 분석 결과 가시화 ( 고빈도 단어, 긍부정 분류 )
    - iii) 오디오 분석 결과 가시화 ( 볼륨의 변화량 )
    - iv) 평준화 오디오 파일 다운로드

#### 백엔드

- 1) 오디오 파일 다운로드 모듈
  - a) 오디오 파일 다운로드 모듈의 입력은 스트리밍 동영상의 주소이며 출력은 mp3 형식의 Audio 파일이다.
  - b) 입력된 url을 youtube-dl 라이브러리의 download 함수의 인자로 넘겨 mp3 형식의 Audio파일을 다운로드
- 2) 오디오 파일 분석 모듈
  - a) 오디오 파일 분석 모듈의 입력은 Video 또는 Audio 파일이며, 출력은 단위 시간당 오디오 평균 볼륨 크기와 평준화된 오디오 파일이다.
  - b) Moviepy 라이브러리를 이용하여 Input 영상 혹은 음성에 대해서 필요한 음성 데이터를 연산이 가능한 배열 형태로 추출
  - c) 해당 배열을 이용하여 단위시간(분) 당 볼륨을 추출
  - d) 영상의 볼륨 평균을 구하여 볼륨값들을 해당값으로 normalize시키고 이를 프론트엔드에서 다운로드 할 수 있게 파일로 write
- 3) 채팅 로그 다운로드 모듈
  - a) 채팅 로그 다운로드 모듈의 입력은 스트리밍 동영상의 주소이며, 출력은 채팅로그(시간, 유저 아이디, 채팅내용)가 담긴 배열이다.
  - b) 입력된 주소가 아프리카TV, 트위치, 유튜브 중 어느 플랫폼인지 판별
  - c) 판별된 플랫폼의 채팅로그를 추출하는 함수를 호출하여 채팅로그 추출
    - i) 아프리카TV - BeautifulSoup을 이용하여 입력된 url을 분석하여 동영상의 고유한 키값을 찾은 뒤 이 값을 아프리카TV의 채팅로그가 저장된 사이트에 params 값으로 넘겨주어 채팅로그를 추출
    - ii) 트위치 - 트위치에서 제공하는 API를 사용하기 위해 트위치 개발자 사이트에서 부여받은 Client ID와 입력된 url을 requests의 get 함수인자로 넘겨주어 채팅로그를 추출
    - iii) 유튜브 - BeautifulSoup을 이용하여 입력된 url을 분석하여 채팅이 표시되는 iframe을 찾아 다음에 추적할 url을 얻어낸다. 이를 request의 get함수인자로 넘겨주어 채팅로그를 추출한다.

 <b>국민대학교 컴퓨터공학부 캡스톤 디자인 I</b>	<b>계획서</b>		
	<b>프로젝트 명</b>	인공지능을 이용한 하이라이트 추천 시스템	
	<b>팀 명</b>	안진마	
	Confidential Restricted	Version 2.3	2020-APR-23

#### 4) 채팅 분석 모듈


- a) 채팅 분석 모듈의 입력은 채팅 로그 다운로드 모듈의 출력인 채팅로그가 담긴 배열이며 출력은 각 함수는 분단위 채팅량을 담은 배열, 채팅에 참여한 총 유저수 또는 채팅에서 많이 출연한 키워드를 담은 배열을 출력한다.
  - i) 채팅량 분석 함수 - 입력받은 채팅 로그에 담긴 채팅 시간을 이용하여 분단위 채팅량을 계산한 뒤 이를 내림차순으로 정렬한 배열을 출력
  - ii) 채팅에 참여한 총 유저수 계산 함수 - 입력받은 채팅 로그에 담긴 유저 아이디를 카운트하여 이를 출력
  - iii) 고빈도 키워드 분석 함수 - KoNLPy 라이브러리의 Okt 형태소 분석기를 이용하여 입력받은 채팅 로그에서 명사를 추출한 뒤 이를 딕셔너리화 하여 출연 빈도를 카운트 한 뒤 출연빈도를 기준으로 내림차순으로 정렬한 배열을 출력

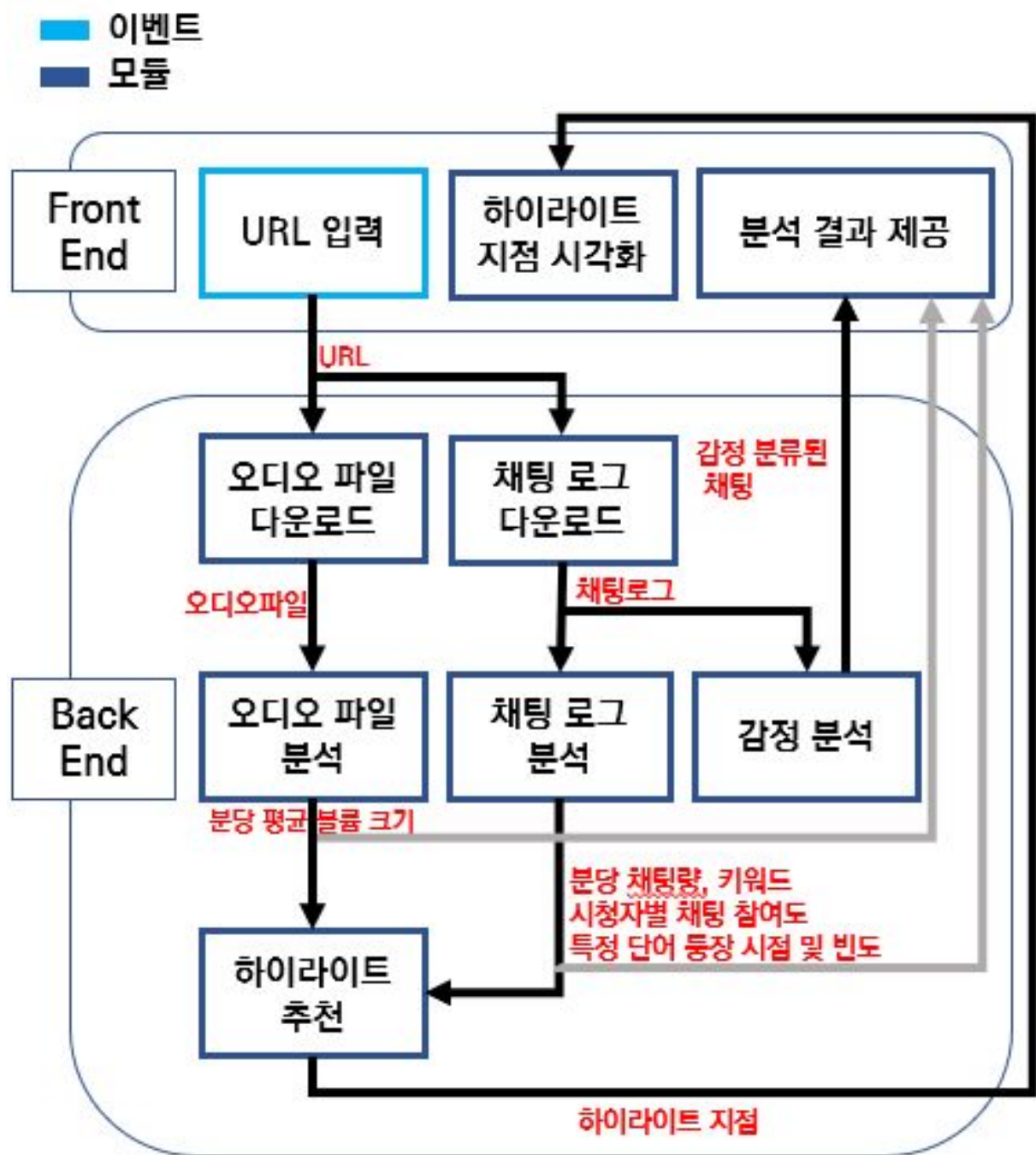
#### 5) 감정 분석 모듈

- a) 네이버 영화 리뷰 데이터셋 150,000개와 라벨링 된 채팅 데이터 15,000개를 사용하여 감정 분석 모델을 학습
- b) 감정 분석 모델은 keras에서 제공하는 Sequential 함수를 이용해 구성
- c) 모델에 말뭉치를 입력 데이터로 주기 위해서는 토큰화 한 뒤 벡터화 해야하는데, 토큰나이저로는 SPM을 사용했고 벡터화 기법으로는 Count Vectorizaion 기법을 사용했다.
- d) 감정 분석 모듈의 Workflow는 다음과 같다.
  - i) 데이터셋 로드
  - ii) SPM 모델 정의(모델 타입은 word, 어휘사전 크기는 10000으로 설정)
  - iii) 데이터 전처리(Tokenize, Vectorize)
  - iv) 모델 훈련
  - v) 모델 평가 및 실제 데이터 추론
- e) 감정 분석 모듈의 입력은 채팅 내용이 들어있는 txt 파일이며, 출력은 긍정/부정으로 분류된 채팅이다.


#### 6) 하이라이트 추천 모듈

- a) 오디오, 채팅 분석 모듈의 출력인 분석 결과를 받아 하이라이트를 추천한다.
- b) 하이라이트 추천 알고리즘은 다음과 같다
  - i) 오디오 분석 모듈에서 볼륨의 크기가 큰 상위 3개 지점을 추출
  - ii) 채팅 분석 모듈에서 채팅량이 많은 상위 3개 지점을 추출
  - iii) 두 모듈에서 리턴받은 결과를 중복 제거 등 후처리 한 뒤 하이라이트 지점을 시:분:초 형식으로 출력
- c) 하이라이트 추천 모듈의 입력은 단위 시간당 오디오 평균 볼륨 크기, 단위 시간당 채팅수이며, 출력은 하이라이트 추천 지점(시간대)이다.

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	인공지능을 이용한 하이라이트 추천 시스템	
	팀 명	안진마	
	Confidential Restricted	Version 2.3	2020-APR-23



[그림 2-8] 시스템 설계도

 <b>국민대학교 컴퓨터공학부 캡스톤 디자인 I</b>	<b>계획서</b>		
	<b>프로젝트 명</b>	인공지능을 이용한 하이라이트 추천 시스템	
	<b>팀 명</b>	안진마	
	Confidential Restricted	Version 2.3	2020-APR-23

### 2.3.4 결과물 목록 및 상세 사양


대분류	소분류	기능	형식	비고
UI	다시보기 동영상 URL입력	동영상을 불러올 URL을 입력한다.	메소드	
	표 및 그래프 보기	동영상 분석이 끝난 후 결과물을 보여준다.	메소드	
알고리즘	채팅로그분석	채팅로그를 분석해 여러가지 결과를 추출한다.	메소드	
	음성 스펙트럼	동영상에서 음성파일을 추출해 스펙트럼으로 분석한다.	메소드	
인공지능	채팅로그 감정 분석	채팅로그를 긍정 / 부정으로 감정을 분석한다.	모델	
DB	영상별 분석 내용	분석결과를 저장한다.	DATABASE	

[표 2-8] 프로젝트 결과물 목록

## 2.4 기대효과 및 활용방안

스마트폰의 발달로 영상을 찍는 일은 간편한 일이 되었다. 하지만 단순히 나를 위한 보관용이 아니라 누군가에게 보여지길 원하는 영상이라면 영상 편집, 콘텐츠 기획등 추가적인 작업이 요구된다. 이 시스템은 크리에이터와 편집자들의 창작활동을 보조하는 수단으로서 다양한 기대효과와 활용방안이 있다.

1. 편집의 지표가 되는 하이라이트 지점 추천과 가시화된 정보를 제공함으로써 편집에 투자하는 시간을 감소시킨다. 기존의 편집은 2시간이 넘는 스트리밍 영상을 여러차례 돌려보며 시청자의 반응을 살피고 이를 기반으로 편집 지점을 찾는 과정으로 진행되었다. 이러한 과정은 매우 긴 시간이 소요된다. 이 시스템은 크리에이터의 영상 분석 데이터를 기반으로 하이라이트 지점을 추천하고 수치화된 정보를 가시화하여 보여줌으로써 편집에 소요되었던 시간을 획기적으로 줄일 수 있다.
2. 긴 편집시간으로 인해 편집자를 고용했던 크리에이터가 자체적으로 편집함으로써 편집자에게 지출되는 비용을 절감시킨다. 영상을 편집하는 과정에서 편집 지점을 모색하는 일은 긴 시간을 투자해야 하는 일이다. 이러한 이유로 크리에이터는 편집자를 고용한다. 이 시스템은 편집지점을 모색하는 시간을 줄여줌으로써 크리에이터는 편집자 고용의 필요성을 감소시키고 이는 편집자에게 지출하는 비용을 절감시킬 수 있다.
3. 크리에이터의 행동에 따른 시청자의 반응 정보를 제공함으로써 크리에이터가 다음 영상 콘텐츠를 기획하는 지표로 사용할 수 있다. 시청자의 반응을 바탕으로 시청자가 보고 싶어하는

 <b>국민대학교 컴퓨터공학부 캡스톤 디자인 I</b>	<b>계획서</b>		
	<b>프로젝트 명</b>	인공지능을 이용한 하이라이트 추천 시스템	
	<b>팀 명</b>	안진마	
	Confidential Restricted	Version 2.3	2020-APR-23

영상을 기획하는 것은 중요하다. 이러한 기획은 양질의 콘텐츠를 생산할 수 있으며 시청자수 증가로 이어진다. 이 시스템은 크리에이터의 영상 분석 데이터를 기반으로 시청자들의 반응 정보를 제공함으로써 크리에이터의 콘텐츠 기획을 돕는다.

- 위와 같은 크리에이터들의 창작활동을 보조하는 기능을 수행함으로써 크리에이터 도전 장벽을 낮추고 크리에이터의 지속적인 성장을 돕는다.

## 3 배경 기술

### 3.1 기술적 요구사항

#### 3.1.1 개발 환경

- 1) 운영체제: Mac OS, Windows 10, CentOS
- 2) 개발 언어: Python 3.7.x, Javascript
- 3) 프레임워크: Flask, Pytorch
- 4) 라이브러리: moviepy, requests, json, BeautifulSoup, numpy, matplotlib, ElementTree, React.js, youtube\_dl, KoNLPy

#### 3.1.2 서버 환경

- 1) 서버: AWS EC2(Nginx)
- 2) 데이터베이스: AWS RDS(Postgresql(DB), Alembic(Migration), SQLAlchemy(ORM))


#### 3.1.3 결과물 확인 환경

- 1) 운영체제: 무관
- 2) 브라우저: Chrome

### 3.2 현실적 제한 요소 및 그 해결 방안

#### 3.2.1 하드웨어

- 1) 하이라이트를 분류하는 모델을 학습시키기 위해 많은 computing power가 요구된다. 모델의 학습은 프로젝트 결과물의 성능과도 직결된 문제로, 이를 해결하기 위해 과에서 제공하는 KMUSW 딥러닝 클라우드를 활용할 계획이다.


 <b>국민대학교 컴퓨터공학부 캡스톤 디자인 I</b>	<b>계획서</b>		
	<b>프로젝트 명</b>	인공지능을 이용한 하이라이트 추천 시스템	
	<b>팀 명</b>	안진마	
	Confidential Restricted	Version 2.3	2020-APR-23

### 3.2.2 소프트웨어

- 1) 외부 시스템에 존재하는 모든 데이터를 미리 서버에 가지고 있는것은 불가능하다. 따라서 하이라이트를 분류하기 위해 영상을 분석하는 시점은 사용자가 URL을 요청한 직후이다. 더불어 모델의 추론 시간까지 포함해야 하므로 실행 시간이 많이 소요될 예정이고, 이것을 해결하기 위해 스트리밍 플랫폼에서 제공하는 개발자용 API를 사용하여 서버로부터 데이터를 직접 받아 사용 할 계획이다.

### 3.2.3 기타


- 1) 의도된 바와 달리 스트리머의 영상이 아닌 경우, 하이라이트 지점을 제대로 분류하지 못할 수 있다. 이는 정규식을 이용하여 다시보기 영상의 URL을 받을 때 잘못 된 경로를 입력 할 시 에러 메시지를 출력하여 해결할 계획이다.

 <b>국민대학교 컴퓨터공학부 캡스톤 디자인 I</b>	<b>계획서</b>		
	<b>프로젝트 명</b>	인공지능을 이용한 하이라이트 추천 시스템	
	<b>팀 명</b>	안진마	
	Confidential Restricted	Version 2.3	2020-APR-23

## 4 프로젝트 팀 구성 및 역할 분담

이름	역할
신상훈	<ul style="list-style-type: none"> <li>- 프로젝트 매니저</li> <li>- 프론트엔드</li> <li>- 데이터 라벨링</li> </ul>
송성재	<ul style="list-style-type: none"> <li>- 오디오 분석</li> <li>- 감성분석 모델 설계 및 학습</li> <li>- 데이터 라벨링</li> </ul>
김연수	<ul style="list-style-type: none"> <li>- 백엔드 구축 및 API 개발</li> <li>- 데이터 라벨링</li> </ul>
박형준	<ul style="list-style-type: none"> <li>- 오디오 분석</li> <li>- 통합 테스트</li> <li>- 데이터 라벨링</li> </ul>
허진선	<ul style="list-style-type: none"> <li>- 데이터 수집</li> <li>- 채팅 분석</li> <li>- 데이터 라벨링</li> </ul>
윤정연	<ul style="list-style-type: none"> <li>- 데이터 수집</li> <li>- 유닛 테스트</li> <li>- 데이터 라벨링</li> </ul>

[표 4-1] 팀 구성 및 역할 분담


 <b>국민대학교 컴퓨터공학부 캡스톤 디자인 I</b>	<b>계획서</b>		
	<b>프로젝트 명</b>	인공지능을 이용한 하이라이트 추천 시스템	
	<b>팀 명</b>	안진마	
	Confidential Restricted	Version 2.3	2020-APR-23

## 5 프로젝트 비용

항목	예상치 (MD)
데이터 수집	3
데이터 라벨링	30
인공지능 설계	70
피쳐 추출	10
프론트엔드 개발	20
백엔드 구축	10
백엔드 API 개발	30
데이터베이스 설계	5
데이터 분석 및 가시화	10
기능 개선	70
유닛 테스트	20
버그 픽스	50
<b>합</b>	<b>328</b>

[표 5-1] 프로젝트 비용




 <b>국민대학교 컴퓨터공학부 캡스톤 디자인 I</b>	<b>계획서</b>		
	<b>프로젝트 명</b>	인공지능을 이용한 하이라이트 추천 시스템	
	<b>팀 명</b>	안진마	
	Confidential Restricted	Version 2.3	2020-APR-23

## 6 개발 일정 및 자원 관리

### 6.1 개발 일정


항목	세부내용	1월	2월	3월	4월	5월	6월	비고
아이디어	문제점 분석							
	솔루션 도출							
	요구 사항 분석							
관련분야연구	기술 조사 및 연구							
	관련 시스템 분석							
설계	시스템 설계							
구현	코딩 및 모듈 테스트							
	시스템 통합							
테스트	시스템 테스트							
	버그 픽스							
	기능 개선							

[표 6-1] 개발 일정

 <b>국민대학교 컴퓨터공학부 캡스톤 디자인 I</b>	<b>계획서</b>		
	<b>프로젝트 명</b>	인공지능을 이용한 하이라이트 추천 시스템	
	<b>팀 명</b>	안진마	
	Confidential Restricted	Version 2.3	2020-APR-23


## 6.2 일정별 주요 산출물

마일스톤	개요	시작일	종료일
계획서 발표	문제점 분석 솔루션 도출 아이디어 구체화 기술 조사 및 연구 <b>산출물 :</b> 1. 프로젝트 수행 계획서 2. 프로젝트 발표 PPT, 영상	2020-01-01	2020-03-27
설계 완료	시스템 설계 프론트엔드&백엔드 스택 조사 <b>산출물 :</b> 1. 시스템 설계도	2020-03-01	2020-03-27
사전 준비	데이터 수집 및 라벨링 서버 구축 데이터베이스 설계 UI/UX 디자인 프론트엔드 UI 구현 <b>산출물 :</b> 1. 라벨링 된 데이터 2. 구축된 서버 3. UI/UX 초안 4. 웹 애플리케이션 프로토타입 v.1.0 (UI/UX, Mock data)	2020-03-27	2020-04-12
중간 자문 평가	인공지능 설계 프론트엔드 개발 데이터 분석 및 통계 백엔드 API개발 <b>산출물 :</b> 1. 웹 애플리케이션 프로토타입 v.2.0 (백엔드 API 호출) 2. 하이라이트 분류 알고리즘 3. 프로젝트 중간 보고서 4. 데이터 시각화 대시보드	2020-04-12	2020-04-23
구현 완료	하이라이트 분류 알고리즘 개선	2020-04-23	2020-05-02

 <b>국민대학교 컴퓨터공학부 캡스톤 디자인 I</b>	<b>계획서</b>		
	<b>프로젝트 명</b>	인공지능을 이용한 하이라이트 추천 시스템	
	<b>팀 명</b>	안진마	
	Confidential Restricted	Version 2.3	2020-APR-23

	프론트엔드 개발 백엔드 개발 시스템 통합 <b>산출물:</b> <ol style="list-style-type: none"> <li>1. 통합된 시스템</li> <li>2. 정확도가 높은 감정분석 인공지능</li> </ol>		
테스트	시스템 통합 테스트 유닛테스트 및 버그 픽스 기능개선 <b>산출물:</b> <ol style="list-style-type: none"> <li>1. 개선된 시스템</li> </ol>	2020-05-02	2020-06-05
최종 보고서	최종 보고 <b>산출물:</b> <ol style="list-style-type: none"> <li>1. 최종 보고서</li> <li>2. 전시용 자료</li> <li>3. 온라인 평가용 자료</li> <li>4. 최종 발표 자료 및 시연 동영상</li> </ol>	2020-06-05	2020-06-12

[표 6-2] 일정별 과업 및 주요 산출물

 <b>국민대학교 컴퓨터공학부 캡스톤 디자인 I</b>	<b>계획서</b>		
	<b>프로젝트 명</b>	인공지능을 이용한 하이라이트 추천 시스템	
	<b>팀 명</b>	안진마	
	Confidential Restricted	Version 2.3	2020-APR-23

### 6.3 인력자원 투입계획


이름	개발항목	시작일	종료일	총개발일(MD)
허진선 윤정연	다시보기 데이터 수집	2020-03-27	2020-03-29	3
전원	채팅로그를 긍정/부정으로 라벨링	2020-03-29	2020-04-13	30
김연수	백엔드 구축	2020-04-02	2020-04-12	10
신상훈	데이터베이스 설계	2020-04-02	2020-04-12	5
송성재 박형준	인공지능 설계	2020-04-02	2020-05-01	70
신상훈 허진선	프론트엔드 개발	2020-04-02	2020-05-02	20
박형준	데이터 분석 및 가시화	2020-04-12	2020-05-01	10
김연수 송성재 윤정연	백엔드 API 개발	2020-04-12	2020-05-02	30
전원	기능 개선	2020-05-02	2020-05-31	70
전원	유닛 테스트 및 버그 픽스	2020-05-02	2020-06-05	70

[표 6-3] 인적자원 투입 계획

### 6.4 비 인적자원 투입계획


항목	Provider	시작일	종료일	Required Options
AWS	Amazon	2020-03-28	2020-06-12	
개인 노트북 6대	개인 소유	2020-03-01	2020-06-12	

[표 6-4] 비 인적자원 투입 계획

 <b>국민대학교 컴퓨터공학부 캡스톤 디자인 I</b>	<b>계획서</b>		
	<b>프로젝트 명</b>	인공지능을 이용한 하이라이트 추천 시스템	
	<b>팀 명</b>	안진마	
	Confidential Restricted	Version 2.3	2020-APR-23

## 7 참고 문헌

번호	종류	제목 및 출처	발행년도	저자	비고
1	웹페이지	Nox Influencer			유튜브 수익 통계
		<a href="https://kr.noxinfluencer.com/">https://kr.noxinfluencer.com/</a>			
2	기사	취미가 돈이 되는 시대... '1인 미디어'가 세상을 바꾸다	2019.08.26	조선희	시청자 선호 플랫폼 통계
		<a href="http://www.donga.com/news/article/all/20190824/97096693/1">http://www.donga.com/news/article/all/20190824/97096693/1</a>			
3	논문	You Only Look Once: Unified, Real-Time Object Detection	2016	Joseph Redmon	객체 탐지 모델
		<a href="https://www.cv-foundation.org/openaccess/content_cvpr_2016/papers/Redmon_You_Only_Look_CVPR_2016_paper.pdf">https://www.cv-foundation.org/openaccess/content_cvpr_2016/papers/Redmon_You_Only_Look_CVPR_2016_paper.pdf</a>			
4	영상	초보 유튜버들이 생각하는 유튜브 고충 ft.수익 공개	2019.11.24	KOQA	인터뷰
		<a href="https://youtu.be/wbsNLMsJm00">https://youtu.be/wbsNLMsJm00</a>			
5	도서	유튜버가 말하는 유튜버: 16명의 유튜버들이 솔직하게 털어놓은 유튜버의 세계	2019.12.16	런업 (김찬준)	인터뷰
		<a href="https://www.bookie.co.kr/book/9788960517592">https://www.bookie.co.kr/book/9788960517592</a>			
6	웹페이지	Overview: module code - MoviePy 1.0.2 documentation		Zulko	Moviepy Docs
		<a href="https://zulko.github.io/moviepy/modules/index.html">https://zulko.github.io/moviepy/modules/index.html</a>			
7	웹페이지	How can I get the volume of sound of a video in Python using moviepy?	2015.01.23	Giannis Iordanou	Moviepy를 활용한 볼륨 추출
		<a href="https://stackoverflow.com/questions/28119082/how-can-i-get-the-volume-of-sound-of-a-video-in-python-using-moviepy">https://stackoverflow.com/questions/28119082/how-can-i-get-the-volume-of-sound-of-a-video-in-python-using-moviepy</a>			
8	깃허브	Command-line program to download videos from YouTube.com and other video sites		sergey.M	youtube-dl

 <b>국민대학교 컴퓨터공학부 캡스톤 디자인 I</b>	<b>계획서</b>		
	<b>프로젝트 명</b>	인공지능을 이용한 하이라이트 추천 시스템	
	<b>팀 명</b>	안진마	
	Confidential Restricted	Version 2.3	2020-APR-23

		<a href="https://github.com/ytdl-org/youtube-dl">https://github.com/ytdl-org/youtube-dl</a>			
9	웹페이지	Getting started with MoviePy - MoviePy 1.0.2 documentation		Zulko	Moviepy tutorial
		<a href="https://zulko.github.io/moviepy/getting_started/getting_started.html">https://zulko.github.io/moviepy/getting_started/getting_started.html</a>			
10	블로그	[Keras] KoNLPy를 이용한 한국어 영화 리뷰 감정 분석		cyc1am3n	네이버 영화 리뷰 데이터셋을 이용한 한국어 텍스트 긍정 부정 분류
		<a href="https://cyc1am3n.github.io/2018/11/10/classifying_korean_movie_review.html">https://cyc1am3n.github.io/2018/11/10/classifying_korean_movie_review.html</a>			
11	깃허브	Unsupervised text tokenizer for Neural Network-based text generation.		google	SPM 토큰나이저
		<a href="https://github.com/google/sentencepiece">https://github.com/google/sentencepiece</a>			

[표 7-1] 참고 문헌