1)

안녕하세요, 저희는 ‘유튜브 비디오 정보를 활용한  
낚시성 콘텐츠 탐지 서비스’ 아이디어로 프로젝트를 진행하고자 하는 강성권, 고은경, 권지혜, 배형준이고 저는 발표자 강성권입니다.

지금부터 7조 아이디어 발표 시작하겠습니다.

2)

발표는 다음과 같은 순서로 진행하겠습니다.

먼저 저희가 아이디어를 도출한 배경에 대해 설명드린 후, 현존하는 유사 서비스와의 차별점을 말씀드리겠습니다. 이후 데이터 처리방안과 분석기법에 대해 다룬 후 마지막으로 예상 활용방안 및 기대효과로 발표를 마무리 하도록 하겠습니다.

3)

낚시성 유튜브 영상으로 인해 시간낭비하신 분들. 많으실 거라고 생각합니다.

4)

실제로 온라인 설문조사 결과, 가장 유해하다고 생각하는 콘텐츠 유형 중 낚시성 컨텐츠에 대한 사항이 상위권을 차지하고 있었으며,

유튜브의 사용량이 증가함에 따라 유튜브 플랫폼 내에서 자극적인 제목을 통한 가짜 컨텐츠 또한 타 매체보다 빠르게 확산되고 있습니다.

한국언론진흥재단에서 실시한 온라인 설문조사 결과, 가장 유해하다고 생각하는 콘텐츠 유형 중 낚시성 컨텐츠가 상위권을 차지하고 있습니다.

유튜브 사용량이 증가함에 따라 유튜브 내에서 자극적인 제목을 통한 가짜 컨텐츠 또한 타 플랫폼보다 빠르게 생산되고 있습니다.

5)

낚시성 컨텐츠를 활용해 영상 클릭량을 올리고자 하는 유튜버들에 대한 문제점을 제기한 뉴스 기사가 다수 포착되었으며(???)

낚시성 영상을 통해 조회수를 올리는 유튜버 문제를 다룬 기사를 확인할 수 있었고

6)

이에 공감하며 일명 ‘기생 튜버’를 비난하는 댓글 또한 확인할 수 있었습니다.

7)

따라서 유튜브 사용자들이 올바른 정보를 습득하고 건전한 플랫폼을 누리는 것을 돕기 위해 저희는 낚시성 컨텐츠를 탐지하는 서비스를 구축할 필요가 있다고 생각하였습니다. 사용자들이 잘못된 내용의 썸네일이나 제목에 휘둘려 시간을 낭비하지 않도록 하는 것이 저희의 목표입니다.

따라서 영상 내용과 연관성이 떨어지는 썸네일이나 제목에 휘둘려 시간을 낭비하지 않는 것에 대한 수요가 있다고 판단하여 낚시성 영상 컨텐츠를 탐지하는 서비스를 기획하였습니다.

아래의 예시를 보면, 제목에선 판빙빙의 근황에 초점을 맞췄지만 실제 영상 내용에선 판빙빙의 근황보다 이중계약, 탈세의혹, 다른 여배우의 실종 사건 등 다른 내용이 더 많습니다. 이런식으로 제목과 내용의 연관성이 떨어지는 영상을 필터링하는 것이 저희 서비스의 목표입니다.

8)

유사한 서비스로 낚시성 뉴스를 판별하는 시스템이 개발되어 있었습니다.

유사한 서비스로 낚시성 기사를 판별하는 시스템이 이미 개발되어 있었습니다.

9)

낚시성 기사 판독기는 저희가 개발하고자 하는 서비스와 분석적인 측면은 같을 것으로 예상되나, 저희의 낚시성 영상 판독기는 기사가 아니라 유튜브 영상을 다룬다는 점에서 도메인의 차이가 있습니다.

현재 유튜브는 타 매체보다 이용 시간 면에서 사용량이 월등히 높으며, 꾸준히 증가하고 있는 추세입니다. 이에 따라 뉴스기사 또한 타 매체보다 유튜브를 통해 접하는 사용자가 급증하여 유튜브 영상 분석의 필요성은 더욱 더 대두될 것입니다.

뿐만 아니라, 요즘 트렌드는 텍스트보다 비디오를 더 많이 접하는 양상을 보인다는 점에서 타 매체의 기사 판독보다 유튜브 영상을 판독하는 것이 더 의미 있을 것으로 보이며 자막, 댓글 등 기사보다 더 많은 텍스트를 다뤄야 하기 때문에 분석으로 얻을 정보가 더 많을 수 있습니다.

뿐만 아니라, 텍스트에서 비디오로 컨텐츠 소비 트렌트가 바뀐다는 점에서 기사 판독보다 유튜브 영상을 판독하는 것이 더 의미 있을 것이라고 생각합니다. 자막, 댓글 등 기사보다 더 많은 텍스트를 다뤄야 하기 때문에 분석으로 얻을 수 있는 정보가 더 많습니다.

10)

또한, 전 세대를 아울러 유튜브가 소비되고 있어 유튜브 영상 분석 시 다양한 세대의 데이터를 얻을 수 있다는 장점이 있습니다.

또한, 전 세대를 걸쳐 유튜브가 소비되고 있어 유튜브 영상 분석 시 다양한 세대의 데이터를 얻을 수 있다는 장점이 있습니다. (영상 데이터? 댓글 데이터?)

11)

다음은 데이터 처리 방안과 분석기법에 대해 설명하겠습니다.

우선 외부적으로 확인할 수 있는 썸네일과 제목을 수집하고, 영상과 영상 내 자동생성자막과 영상 댓글을 크롤링을 통해 수집합니다.

다음으로 수집한 데이터들을 다양한 방법을 통해 데이터 전처리를 시행합니다.

우선 썸네일의 경우, 썸네일 내 텍스트와 이미지의 텍스트 detection과 object detectio을 동시에 시행하여 label들을 추출합니다.

우선 썸네일의 경우, 썸네일 안의 글자와 사물을 탐지하여 label을 추출합니다.

제목의 경우 text tokenization을 진행하고 정제합니다.

제목의 경우 text tokenization을 이용하여 정제합니다.

다음으로 영상에 대한 summarization을 진행합니다. Seq2seq라는 모델을 통해 영상 내 자동생성자막을 요약한 뒤, text tokenization을 시행합니다.

마지막으로 댓글의 감성분석을 시행하여 댓글이 부정적인지 긍정적인지 분류합니다.

12)

데이터의 전처리가 끝나면, 분석을 진행합니다.

기본적으로는, 썸네일, 제목, 동영상 정보, 영상 Summarization 4가지 정보들간의 일관성을 비교할 것입니다.

이와 더불어 댓글, 좋아요/싫어요 변수를 통해 낚시성 콘텐츠 탐지해볼 수도 있을 것 같습니다.

최종적으로 컨텐츠가 낚시 영상일 확률을 산출할 수 있습니다.

13)

서비스 제공 방안으로는 다음과 같이 2가지를 생각하고 있습니다.

해당 영상의 링크를 입력하면 낚시성 영상일 확률이 산출되게 하는 웹페이지를 구현하는 것입니다. 낚시성 영상일 확률 이외에도 해당 영상 댓글의 긍정/부정 비율 등도 함께 확인할 수 있습니다.

그리고 만약 가능하다면, 이 서비스를 크롬 확장 프로그램으로 구현하여 마우스를 영상 썸네일 위에 올리면 해당 영상이 낚시성 컨텐츠일 확률을 표시되게 할 수 있습니다.

14)

저희 서비스를 통해 일반 유튜버는 보다 양질의 컨텐츠를 생산하기 위한 기준을 마련할 수 있으며,

유튜브 채널을 운영하며 마케팅에 활용하고자 하는 기업은 낚시성 콘텐츠 제작으로 인한 기업 충성도 하락을 방지할 수 있습니다.

영상 콘텐츠를 소비하는 소비자들은 정보의 홍수 속에서 양질의 정보를 선택하는 것이 용이해지며, 저희 서비스를 이용하는 이용자들이 많아짐에 따라 사회에서는 낚시성 콘텐츠로 인한 정치적 혼란 방지가 가능합니다.

(서비스의 메인 수요자가 소비자라는 것을 언급하고 싶은데 어떻게 해야 할 지 모르겠네요.)

그럼 여기에서 낚시성 컨텐츠 탐지 서비스와 관련된 7조 아이디어 발표를 마무리하겠습니다. 경청해주셔서 감사합니다.