5월 14일 회의록

* 유튜브 최대 크롤링 개수 제한있음
* >> 유튜브의 채널의 특성을 분류하자

공영 vs 케이블

공영 - kbs 뉴스, mbs 뉴스

-> 각자 두개 정도

정치나 어느 회사 밑에 있는지 등의 정보 파악

해당 채널 관련해서 의미있는 정보는 가져오기

현재 댓글의 경우, 한방향으로 치우쳐져 있는 상황

-> 정치 관련 여론 조사에서는 추세와 경향(찬반)을 보는데

추세의 경우 의 비율을 확인하기

-> 댓글의 시간에 따라 달리는 지 확인해야하는 지(시간별 긍정 부정)

댓글 전처리 문제

-> 데이터 수집 후 쪼개서 본 후 감정 분석을 할 수 있다면 하고,

나중에 댓글을 가져온 후 국어 문법에 의거하여 조사를 빼고 등등 구분을 하는데,

이러한 전처리 과정이 어려울 것이라 예상전 단위가 뒷 단위에 주는 영향이 있다. 시간이 갈 수록 줄어듦

요일 질문 사항

* SQL 로 추출 ? 뽑아오기 SQL 자체로도 개발이 가능? DB연동?

개인 과제

* 채널 사전조사

차후 과제

* 네이버 댓글 수집은 차후 과제로