**유튜브 댓글분석 및 정보 검출**

**정영환**

**<분석 방향>**

유튜버들에게는 댓글이 굉장히 많이 달림

많은 방문자들을 나누어보면

1. 우연찮게 누른사람 (열심히 보지도 않고 관심도 없음)
2. 몇번 보고는 있지만 인기 있는 영상한두개만 인식하는 사람(충성도 낮음)
3. 자주보고 챙겨보는 사람 + 구독 하지만 댓글은 달지 않음 (다수 시청자)
4. 3번의 특징을 포함 하면서 꼬박꼬박 댓글달아 유튜버에게 메시지를 보내는 사람(확률 높음)
5. 4번의 특징을 포함 하면서 댓글을 다는 사람들끼리 카르텔이 형성되어 소통하는사람(골수팬)
6. 5번의 특징을 포함 하면서 해당 물건을 자주 구매 하고 비교 댓글까지 남기는 사람 (거의 확실한 소비자)

수많은 시청자들 중에 우리가 주목할 사람은 4,5,6에 해당하는 댓글 다는 사람에 해당 하고

댓글 다는 사람들 중에서 물건을 구매할 것 같은 고객층을 보유한 사람을 선별하는 것이 댓글 분석의 과제

\*물론 3번처럼 댓글도 안 달고 그냥 보기만 하다가 물건 사버리는 사람도 있고 2번처럼 잠깐 지나가다가 보고 어? 좋네 하고 사는 사람도 있음 그런 사람은 유튜브 구독자 증감 추이나 해당 유튜버의 마케팅 방식에서 따로 모델링하여 합산할 것임

**1단계 핵심 문장 찾기**

-미리 만들어진 충성 고객의 댓글 데이터를 수집 및 분류

Ex) 언니 리뷰 감사해요 저번것보다 이게 더 좋아 보이네요 이거 다쓰면 살게요

**2.단계 새로운 댓글들 중에 우리가 가지고 있는 데이터(리뷰를 보고 산사람 혹은 살것 같은 사람)과 비교 해서 살 것 같은 댓글을 단 문장을 카운트**

Ex> 우리가 가지고 있는 재구매 고객의 대화 형식 :

언니 리뷰 감사해요 저번것보다 이게 더 좋아 보이네요 이거 다쓰면 살게요

새로운 댓글 : 언니 너무 예뻐요 저도 이거 살래요 저번 브러셔 보다는 이게 더 가성비가 좋아 보여요 ㅜㅜ

* **문장의 유사도를 평가하고 핵심 단어의 가중치를 분석해 일치 여부를 판단**

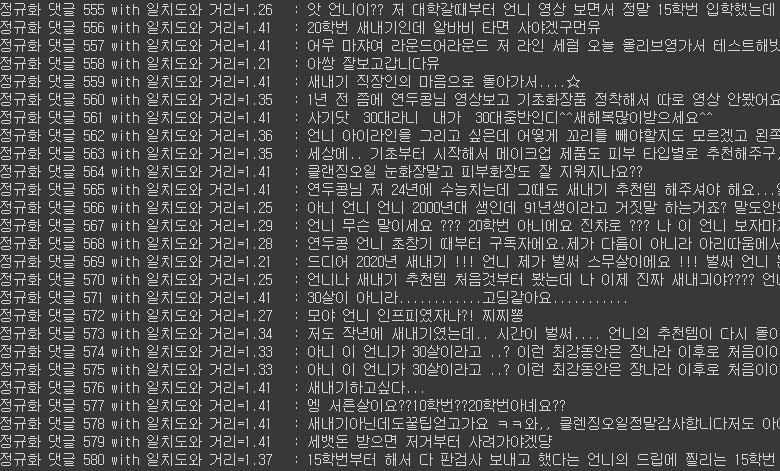
**3단계 해당 문장을 단 유저의 고객층 분석 및 정보 수집**

Ex> 10대 여성, 구매 여부, 재구매 의사 등등…….찾을수 없어도 상관 없음

**4단계 리포트 화 예정**

**현재 1,2,단계 핵심 문장 분류 + 데이터 분석을 통한 문장 카운트 진행중**

**결과>**



**새로운 댓글과의 분석을 통해 비슷한 문장을 찾아 우호도 및 구매의사 평가**

