

유사도 기반의 군집화를 통한 SNS 광고 게시물 필터링

Final Paper Draft

**2017. 5. 24**

|  |  |
| --- | --- |
| 과 목 | TIM 61101 빅데이터와 신제품 개발 |
| 담당교수 | 임동준 교수님 |
| 이름 및 학번 | 차욱현 |
|  | 장진규 |
|  | 이정민 |
|  | 양지은 20168216 |

**목 차**

1. 서론
   1. 프로젝트의 배경 및 목적
2. 분석 내용
   1. 데이터 소스와 분석 환경
3. 분석 결과
4. 결론
5. **서론**
   1. **프로젝트의 배경과 목적**

오늘날 개인적 선호나 취향과 관련된 정보를 검색하거나 접하는데 가장 많이 사용되는 매개체 중 하나는 SNS(Social Network Service)이다. 주변 사람들 및 타인들이 올리는 게시물을 보면서 새로운 레스토랑의 메뉴를 접하고 요즘 뜨는 앱을 알게 되며 여행지 사진을 공유하기도 한다. 특히 과거 블로그의 장황한 글 보다는 이미지나 동영상으로 시각화되어 있기 때문에 빠르고 쉽게 가장 트렌디한 정보를 접하고자 하는 유저에게 선호된다.

SNS를 이용하는 주요한 이유 중 하나는 실제 사용자의 후기나 느낌을 알고 싶기 때문이다. 레스토랑을 찾을 때 또는 어떤 서비스를 이용할 때 우리는 홈페이지보다는 개인이 작성한 리뷰를 선호하며 특히 여행이나 맛집과 관련해서는 블로그, Tripadvisor, 구글맵 리뷰 등이 많이 사용되고 있다. 그 중에서도 최근 몇 년 간 사용자가 급격하게 늘어나고 있는 인스타그램은 말그대로 Social Network를 기반으로 하는 서비스이므로 주변 사람들의 경험이나 느낌을 쉽게 접할 수 있지만 이러한 특성 때문에 상업적 목적을 가진 이용자가 게시글에 자연스럽게 광고를 녹여내기에도 적합하다. 이 프로젝트는 매일 매일 올라오는 새 게시물 사이에서 수많은 광고를 걸러내고 볼 수는 없을까? 라는 의문에서 출발하였다.

따라서 이 프로젝트의 목적은 Instagram에서 올라오는 맛집과 관련된 게시물 중 광고성 게시물을 필터링하여 순수 게시물만 나타내는 것을 구현하는 것이다. 이전 프로젝트인 ‘여기 서봐’가 사용자의 위치와 취향을 고려한 사진을 추천해 주는 것이 목적이었다면 변경된 프로젝트는 사용자의 취향을 분석하여 맛집을 추천해 주는 것이 아니라 어떠한 키워드를 검색했을 때 광고로 예측되는 게시물을 필터링하여 표시되지 않게 하는 것에 중점을 두고 있다.