**비정형 데이터 활용 분석 프로젝트 기획안**

**2024년 5월 21일**

**과정명: 금융&마케팅 데이터 분석(Python)**

|  |  |
| --- | --- |
| **프로젝트 주제** | **강남구 치킨집의 브랜드 평판분석** |
| **팀 명** | 데이터 더 조 |
| **팀 원** | 김현, 백진희, 심성우, 이동욱 |
| **프로젝트 내용**  해결하고자 하는 문제  최종 산출물의 청사진 | 1. 배경 : 네이버 검색 시 상단에 상위 노출되는 매장들은   사용자가 신뢰할 수 있는 근거에 기반한 지표인가?   1. 문제인식 : 네이버에 “강남구 치킨“이라고 검색 할 경우 상위   노출되는 6개 매장은 단순 광고가 아닌, 긍정적인  소비자 평가에 의한 노출 결과인지 확인하고자 합니다.   1. 해결방안 : 네이버 지도에서 강남구 치킨 검색결과 70개 매장에 대한   호감도를 분석하여 순위를 분석하고, 네이버 검색 시  상위 노출되는 6개 강남구 치킨 매장의 호감도 순위와  비교합니다. |
| **프로젝트 추진 일정** | 1. 데이터수집 : 네이버 크롤링05/20 ~22 2. 데이터 전처리 : 필요한 특성 추출05/23~26 3. 데이터 모델링 및 평가 : 05/27~28 4. 발표 준비 : 05/28~30 |
| **프로젝트 수행 방향**  수행 방법/도구 | 1. 수행방법/도구   **<브랜드 평판분석>**  호감도 : 플레이스 리뷰 수, 플레이스 리뷰 감성분석을 통한  호감도산출  네이버 지도에서 (“강남구 치킨”) 검색  -> 네이버 지도 Selenium을 이용한 동적 크롤링 및 감성분석  전처리 : 의미 없는 리뷰 제거, 누락값 처리 등  감성분석 모델 : WhitePeak/bert-base-cased-Korean-sentiment)를  이용한 감성분석 후 호감도 산출  호감도 = 긍정 / 긍정+부정 (단위:건수) \*  \* 출처 : 송은지. (2013). 빅 데이터를 이용한 고객평판 사례분석에 관한 연구. 한국정보통신학회논문지, 17(10), 2439-2446. |