**[ 1차(5월 5주차) 프로젝트 수행 일지 ]**

| **프로젝트 타이틀** | 판매 데이터 분석을 통한 수요 예측 | | |
| --- | --- | --- | --- |
| **프로젝트 팀명** | 3조 - 데이터 맛집 | 프로젝트 팀원 | 황진웅(팀장),  김철기, 박건하, 성원선, 홍석우 |

※ 프로젝트 수행 일지는 매 주 금요일 17시까지 최종본을 저장해주세요

**1. 프로젝트 수행 계획 및 현황**

| **1. 프로젝트 수행 계획 및 현황** | | |
| --- | --- | --- |
|  | **금주 프로젝트 수행 내용** | **차주 프로젝트 수행 계획** |
|  | **1) 주제 선정 회의**  - 판매자 등급화와 가격비교 데이터를 이용한 수요 예측과 재고 관리 최적화  - 지난 해의 클릭량을 기준으로 수요 예측  - 날씨나 감정을 기반으로 트렌드 분석  - 리뷰 데이터를 통해 자동 별점 부여 기능  - 이용 가능한 데이터 파악과 활용 가능한 데이터들에서 인사이트 도출(데이터와 자료를 수집한 뒤 역으로 주제를 선정)  **2) 주제 선정**  - 주제 : 마켓컬리 리뷰 분석 기반 자동 평점 및 맞춤형 상품 추천 시스템  - 확장 :  **(1) 자동 평점 산출 시스템**  - 해시태그-키워드 자동생성  (사용자 후기 탐색 피로도 감소)  **(2) 맞춤형 상품 추천 시스템**  - 다양한 조건과 사용자 취향에 맞춘 정밀 추천 기능(날씨, 시간, 취향 고려)  - 챗봇 기능(사용자의 현재 상황 고려)  **3) 기획안 작성**  - 주제 선정 회의를 기반으로 기획안 초안작성 및 강사님 및 멘토님의 피드백을 바탕으로 기획안을 재구성  **4) 데이터 수집**  - 11번가, 옥션에서 카테고리별 리뷰 데이터 크롤링(학습용)  - 마켓컬리 리뷰 데이터(테스트용)  **5) 제안받은사항**  **[강사님]**  - 특정 카테고리 중심의 모델 학습 후 일반화 가능성 검토 제안  - 타 사이트 또는 카테고리의 리뷰데이터의 학습도 긍정, 부정을 충분히 분류 가능함을 안내  - 감정 기반 상품 추천의 경우, 감정-음식 매핑  - 학습에 필요한 데이터를 적정량 수집 후 빠르고 원활한 데이터 분석 제안(데이터 간의 상관관계와 활용 방안 체크)  **[멘토님]**  - 최종적인 결과를 시각화 해보고 추후 데모 페이지 구현 제언 | **1)크롤링 데이터 결과 확인**  - 개별 수집 데이터 취합 및 공유  - 수집한 리뷰 데이터의 결측치 확인 및 데이터 추가 크롤링  **2) EDA 및 전처리**  2.1) 시각화  - 평점 분포  - 리뷰 길이 분석  - 키워드 빈도수 분석  - 리뷰내용 이모지 빈도수 분석  2.2) 데이터 정제  - 특수문자 및 이모지 정리  - 리뷰 감성 라벨링  - 이상치 제거  **3) 수집한 데이터 분석**  - 샘플 데이터의 학습을 통해 수집한 데이터간의 상관관계 확인  - 이모지, 도움되요 등 추가 피처를 반영하여 모델 성능 변화 비교 분석  **4) 데이터 학습 테스트**  - 전처리와 분석을 마친 데이터를 학습시키기 전 성능테스트 및 회의를 통한 코드공유  - 11번가와 옥션의 리뷰 데이터(리뷰, 평점) 비교 분석을 통해 학습 데이터 선정  - 리뷰 데이터의 긍정 및 부정 분류 확인 |

**2. 강사님 피드백**

| 팀장, 팀원으로 나눠서 간단히 적기보다는 각 인원별 수행 내역이나 앞으로 수행할 방향도 분류해서 적어주시면 좋을 듯 합니다.  흐름은 상세히 나눠서 적어주셨는데 세부 내용들이 구체적이지 않네요. 예를 들어 왜 마켓컬리가 주제인데 11번가, 옥션도 추가적으로 크롤링을 하는지 등이 나와있으면 좋을 것 같습니다. 핵심들이 빠진 것 같네요 ㅎ  오늘 피드백 드린 전처리와 EDA 등이 왜 필요한지 어떻게 처리를 할지 틀려도 좋으니 고민해보시고 작성해보시는 것도 좋을 것 같습니다. 이후에 스스로 피드백이 가능하게 설계할 수 있습니다 ㅎ |
| --- |