**프로젝트 진행일지**

**[스타트 라이트]**

|  |  | | **작성자** | 홍승표 | |
| --- | --- | --- | --- | --- | --- |
|  |  | | **날 짜** | 2023.05.07. (14회) | |
|  |  | | **진척률** |  | |
| **팀 명** | 스타트라이팀 | | | | |
| **팀구성원** | **이름** | **역할** | | | |
| 서단우 | 상권 분석용 데이터 수집, 데이터 분석, 데이터 전처리, 오피니언 마이닝 및 모델 학습, 데이터 시각화 | | | |
| 조은진 | 웹 크롤링, 오피니언 분석용 데이터 수집, 데이터 분석,  오피니언 마이닝, 데이터 시각화(인구 분석) | | | |
| 차지인 | 기획 발표 준비, 오피니언 분석용 데이터 수집, 데이터 분석,데이터 시각화(매출 분석), 언어 모델 비교 | | | |
| 함은정 | 상권 분석용 데이터 수집, 데이터 분석, 데이터 시각화(지역 분석), 최종 발표 자료 제작 | | | |
| 홍승표 | 기획 발표 준비, 오피니언 분석용 데이터 수집, 데이터 분석, 데이터 시각화(업종 분석) | | | |
| **일별 진행상황** | | | | | |
| **이름** | **업 무 내 용** | | | | **특이사항** |
| 서단우 | 오피니언 마이닝 모델 테스트 | | | |  |
| 조은진 | 오피니언 마이닝 모델 테스트 | | | |  |
| 차지인 | 언어 모델 비교 | | | |  |
| 함은정 | 최종 발표 자료 제작 | | | |  |
| 홍승표 | 태블로 시각화 | | | |  |

| **회 의 내 용** |
| --- |
| [안건 1]   * 2차 프로젝트 진행 과정   [세부내용]   * 발표 피피티 제작 어느정도 완료 * 발표에 들어갈 내용 수정 필요  1. 모델 선정 과정  * 회귀 모델 vs 분류 모델 * 회귀모델 : R2 score ~ 0.9 : 상당히 높게 나왔다. * 그러나 RMS 100 대 -> 백만원 단위에서 오차 발생 가능 * 분류 모델 : 정확도가 상대적으로 낮음 (~0.47) * 매출을 8개의 구간으로 비교적 균등하게 잘라서 진행 * 히트맵으로 봤을 때, 최저구간과 최고구간은 맞추지만, 그 가운데 영역은 어려워 하는듯 * 내일 : 발표 예정 |