**[ 전공 프로젝트 1주차 수행 일지 ]**

| **프로젝트 타이틀** | 자영업자 대상 상권 & 타겟 고객 분석 및 추천 모델링 | | |
| --- | --- | --- | --- |
| **프로젝트 팀명** | **우리가게물건사조** | 프로젝트 팀원 | 김준호, 정수연 |

**1. 프로젝트 수행 계획 및 현황**

| **1. 프로젝트 수행 계획 및 현황** | | |
| --- | --- | --- |
| **이름** | **금주 프로젝트 수행 내용** | **차주 프로젝트 수행 계획** |
| **데이터**  **사이언스** | 1. 주제 선정   * 2022 카타르월드컵 데이터 분석, 팀간 승률예측 * **2021 상권 및 고객 분석 & 추천 모델링**   2. 데이터 수집 (Seoul 열린 데이터 광장)   * CSV: 서울시\_우리마을가게\_상권분석 (신\_상권\_추정매출)\_2021년 * OpenAPI: 서울시\_우리마을가게\_상권분석 서비스(신\_상권영역)   3. 데이터 전처리   * CSV, Open API 데이터 결합 * 외식 업종만 선별   (‘서비스\_업종\_코드’ : CS1으로 시작)   * ‘외식업\_연매출\_금액’ 파생변수 생성 * 필요한 columns로 새로운 데이터 프레임 생성   4. 멘토링 및 팀미팅 준비   * 데이터 건수 자료를 토대로 서비스화하기 위해 데이터를 어떤 방식으로 정리하는게 좋을지? * 요일 (주중,주말), 시간, 성별, 연령 4가지 고객 특성을 간략화하는 방식을 어떻게 하면 좋을지? (Max 값으로 통일하는 방안) | 1. 데이터 분석   * 외식업 top 10 상권 vs 세부 외식 (10개) 업종별 top 10 상권 비교 * 세부 외식업 top 10 상권 상세 분석 (요일, 방문시간, 성별, 연령)   2. 데이터 시각화   * 외식업 매출 Top10 상권 시각화(막대) * 세부 외식업(10개) 매출 Top10 상권 시각화(막대) * 지도 내 외식업 TOP100 상권 파악 * 지도 내 세부 외식업(10개) 요일, 방문시간, 성별, 연령별 시각화(파이차트) * 연 매출액과 타 변수와의 상관관계 시각화 |

**2. 강사님 피드백**

| **2. 강사님 피드백** | |
| --- | --- |
| **데이터 사이언스**  **문경미 강사님** | 머신러닝을 위한 데이터의 전처리가 필요 |