|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 참가 서약서 및 개인정보 수집·이용 동의서 | | | | | | | | |
| **참가 서약서** | | | | | | | | |
| 본인은 CJ대한통운이 주최하는 ‘미래기술 챌린지’에 참가함에 있어, 다음 각 호의  규정을 성실히 준수할 것을 서약합니다.  1. 대회 제반 규정을 준수하며, 이를 준수하지 않을 경우 불이익이 발생할 수 있다.  2. 이미 상용화된 작품, 타 대회 출품 수상작, 타인의 저작권을 침해한 작품은 참가  자격 취소 및 입상 무효 처리될 수 있다. | | | | | | | | |
| **개인정보 수집·이용 동의서** | | | | | | | | |
| **가. 개인정보 수집·이용 목적**  1) 「미래기술 챌린지」 운영 및 관리(공지사항 전달, 심사 및 수상작 선정, 증명서 및  상장 발급, 수상작 관리 등)를 위해 필요한 최소한의 개인정보(참가 신청서상  기재된 사항)만을 수집·이용하며, 목적 외의 용도로는 활용되지 않습니다.  **나. 개인정보 수집 항목**  1) 성명, 생년월일, 소속, 직위(학년), 이메일 주소, 휴대폰 번호  **다. 개인정보 보유 및 이용 기간**  1) 개인정보 수집 및 이용 목적 달성 시까지(경진대회/상용화 진행 기간 및 유관 교육,  행사 공지) 보유하며, 개인정보 수집 및 이용목적이 달성되면 지체 없이 파기하는  것을 원칙으로 합니다.  **라. 개인정보 수집·이용 동의를 거부할 권리 및 동의 거부에 따른 불이익**  1 정보주체는「미래기술 챌린지」의 개인정보 수집·이용 동의를 거부할 권리가  있습니다.  2) 대회 운영 및 관리에 필요한 최소한의 개인정보 수집·이용 동의에 거부할 경우  본 대회 참여가 불가능합니다.  본 참가자는 위 참가서약서와 개인정보 수집·이용 동의서의 내용을 모두  확인하였으며 이에 동의합니다.  ■ 동의 □ 미동의  2021년 10월 28일   |  |  |  |  |  | | --- | --- | --- | --- | --- | | 성명 | 소속 | 생년월일 / 성별 | 동의여부 | 서명 | | 강동인 | 서강대학교 | 970416 / 남 | ■ 동의 □ 미동의 |  | | 박 건 | 서강대학교 | 980615 / 남 | ■ 동의 □ 미동의 |  | | 허채원 | 서강대학교 | 990626 / 여 | ■ 동의 □ 미동의 |  | |  |  |  |  |  | | | | | | | | | |
| 기획 아이디어, SW 개발소스 제3자 공개 동의서 | | | | | | | | |
| **가. 제3자 공개 및 공유**  ◦ 본인은 CJ대한통운이 아이디어 및 개발소스를 ｢미래기술 챌린지｣의 취지, 목적을  달성하기 위해 필요한 한도 내에서 복제, 전송의 방법으로 이용하고나 제3자에게  공개 및 공유하는 것에 동의합니다.  - - 응모 아이디어 및 앱에 대한 평가와 대회 관리 및 운영에 관련한 업무 수행  - 전문가가 참여하는 평가 및 심사  - 대회 개최 이후 앱 개발 활성화 및 연구기술개발의 목적 활용  **나. 수집 항목**  ◦ 아이디어, 개발소스  **다. 앱 기획 아이디어 보유･이용기간**  ◦ CJ대한통운은 공모작을 공모전 종료일로부터 6개월 이내에 모두 폐기한다.  단, 폐기 전에 저작자가 저작물의 반환을 요구할 시 저작자에게 반환 하여야 한다.  ◦ CJ대한통운은 입상작에 대한 저작재산권을 합당한 보상을 전제로 우선적으로  양수할 수 있다. 이 때, 양도 여부, 범위 및 그 대가에 대해서는 추후 저작자와  별도 협의를 통하여 정하는 것으로 한다.  **라. 기획 아이디어, 앱, 소스 제3자 공개･공유에 동의하지 않을 권리 및 동의하지 않을 경우의 불이익**  ◦ 정보주체는 ｢미래기술 챌린지｣에 제출한 기획 아이디어, SW 개발소스의 공개․  공유를 거부할 권리가 있습니다.  ◦ 기획 아이디어, 개발소스 공개․공유에 동의하지 않을 경우, 본 공모전 참가신청이  불가 합니다.  본인은 ｢미래기술 챌린지｣에서 본인의 기획 아이디어, 개발소스의 제3자 공개·공유에  동의합니다.  ■ 동의 □ 미동의  2021년 10월 28일   |  |  |  |  |  | | --- | --- | --- | --- | --- | | 성명 | 소속 | 생년월일 / 성별 | 동의여부 | 서명 | | 강동인 | 서강대학교 | 970416 / 남 | ■ 동의 □ 미동의 |  | | 박 건 | 서강대학교 | 980615 / 남 | ■ 동의 □ 미동의 |  | | 허채원 | 서강대학교 | 990626 / 여 | ■ 동의 □ 미동의 |  | |  |  |  |  |  | | | | | | | | | |
| 개발 계획서 | | | | | | | | |
| 주제 | 주말 주문 데이터를 활용한 월요일 주문량 예측 | | | | | | | |
| 팀구성원 | No. | 구분 | 성명 | 소속 | No. | 구분 | 성명 | 소속 |
| 1 | 팀장 | 박건 | 서강대학교 | 2 | 팀원 | 강동인 | 서강대학교 |
| 3 | 팀원 | 허채원 | 서강대학교 |  | | | |
| 1. 개 요 | **1.1 배경 및 필요성**  과거에는 직접 쇼핑몰에 방문하여 상품을 구매하는 오프라인 쇼핑이 주를 이뤘다면, 정보통신 기술의 발달에 따라 온라인 쇼핑의 비중이 점점 커지고 있다. 뿐만 아니라 코로나19로 인해 외출이 어려워지자 온라인 쇼핑몰 사용량은 더욱 증가하고 있다. 더욱 확대되는 온라인 쇼핑 시장과 이에 따른 택배 시장의 변화에 대처하기 위해서는 데이터 기반의 정확한 수요량 예측이 필요하다.  EDA 결과, 각 요일 중 월요일은 주문량이 가장 많은 요일이다. 정확한 주문량 예측은 물량의 효율적 관리 및 소비자 만족도 제고를 위한 가장 중요한 요소이다. 특히 갑작스레 쏟아져 나오는 수요는 제조, 물류, 서비스 전반에 걸쳐 반드시 극복해야 하는 문제이다. 이에 소비자의 구매 활동에 영향을 미치는 여러 요인을 고려하여 주문량 예측에 대한 정확성을 높이고자 한다.  **1.2 목적**  택배 서비스의 효율적 관리 및 품질을 결정하는 기준은 ‘하루’ 단위이다. 따라서 주말 데이터를 활용하여 ‘제품별 월요일 주문량’을 예측하고자 한다.  **1.3 혁신성**  주문량을 예측하는 데 있어 기존에 주어진 변수들뿐 아니라, ‘날짜’와 ‘요일’ 변수 사이에 존재하는 사람들의 인문사회학적, 심리적 특성을 반영하고자 하였다. 예컨대 소비자의 요일별 주문량에는 특정 패턴이 존재하는데, 이는 일주일 중 사람들이 사회활동을 하는 주중과 휴식을 취하는 주말의 활동 양상이 다르기 때문이다. 한편 온라인 쇼핑몰과 대립 관계에 있는 것처럼 보이는 오프라인 쇼핑몰의 사용량은 오히려 온라인 쇼핑몰 주문량을 예측하는 데 영향을 주기도 하는데, 이를 요일별 주문량 패턴과 함께 고려하면 더욱 정확한 예측이 가능하다.  이렇게 직접적으로 드러나진 않지만, 주문량에 주요 영향을 미치는 변수들을 반영하여 참신한 관점으로 주문량 예측을 하고자 하며, 이는 예측의 정확성을 높이는 데 기여할 수 있을 것이다. | | | | | | | |
| 2. 과제  수행 | **2.1 수행 방법**   1. EDA : 각 컬럼 특성 분석, 요일 별 주문량 분석 (허채원) 2. 데이터 엔지니어링 : 모델을 위한 train&test data set 및 파생변수 생성 (박건) 3. 예측 모델 구축 : ‘catboost’ 모델을 활용한 모델링 (강동인)   ‘xgboost’ 모델을 활용한 모델링 (박건)   1. 검증 : 각 모델 rmse 계산, 더 좋은 성능을 보인 catboost 모델 확정.   **2.2 수행 과정**   1. EDA  * ‘BKG\_DATE’ 시계열 데이터 변환, ‘day’라는 새로운 요일 column 생성. * Pivot Table을 통해 제품별 각 요일 주문수량의 총합 도출 * 각 총합 정규화 (주문량 규모 차이에 따라 왜곡 발생을 막기 위함) * 시각화      1. 데이터 엔지니어링   - [‘날짜’, ‘요일’, 'CORP\_ID', 'BKG\_TYP', 'ITEM\_CD', 'ITEM\_QTY', 'ITEM\_AMT', 'DLVCLSFCD', 'POST\_ZONE'] 컬럼 사용.  - 날짜 데이터를 통해 ‘n주차’ 컬럼 생성  - 제품별, n주차별 주문 수량 데이터 생성  - ‘ITEM\_AMT’/’ITEM\_QTY’를 통해 파생변수 ‘price’ 생성  - 데이터 엔지니어링을 통한 최종 data\_set 생성    (target=월요일 주문량, SAT=토요일 주문량, SUN=일요일 주문량)   * Data\_Set으로 Train\_set, Test\_set 분리  1. 예측 모델 구축   -ITEM\_CD, CORP\_ID, BKG\_TYP, DLVCLSFCD, POST\_ZONE 변수 카테고리화  -price 변수 minmax\_scale transform  -iqr을 통한 이상치 제거  -catboost, xgboost 두 개의 모델로 예측 및 rmse 도출   1. 검증(rmse)   -Catboost : 4.52 / XGBRegressor : 4.72  -성능이 더 좋은 Catboost로 모델 확정 | | | | | | | |
| 3. 활용방안 및 기대효과 | **3.1 활용방안**  - 상품 수요를 고려한 재고 및 창고 관리  - 효율적 재고 관리를 통한 배송시간 단축  - 예상 주문량에 따른 유동적 서비스 운영  **3.2 기대효과**  - 주문이 몰리는 월요일의 주문량에 대한 정확한 예측의 기대효과는 CJ대한통운, 소비자 입장에서 나눠 생각해볼 수 있다. 먼저 CJ대한통운은 주말의 주문량을 기반으로 차주 월요일의 주문량을 정확하게 예측함으로써 보다 효율적이고 원활한 물량 관리가 가능해진다. 또한 소비자에게는 정확한 배송 정보, 신속한 배송 서비스 등을 제공함으로써 CJ대한통운에 대한 소비자의 만족도 또한 높아질 수 있다.  - 월요일뿐 아니라 일주일 중 공휴일이 존재하는 경우, 공휴일 데이터를 기반으로 공휴일 직후 날짜에 대한 주문량 또한 예측할 수 있다. | | | | | | | |