

식음업장 메뉴 수요 예측 AI 온라인 해커톤

개요

[배경]

리조트 내 식음업장은 계절, 요일, 투숙객 수, 행사 일정 등 다양한 요인에 따라 수요가 크게 변동하는 환경에 놓여 있습니다. 특히 휴양지 리조트는 단기간에 집중되는 고객 수요와 예측하기 어려운 방문 패턴으로 인해, 메뉴별 식자재 준비, 인력 배치, 재고 관리에 있어 높은 운영 난이도를 가집니다.

이러한 복잡한 운영 환경 속에서 정확한 메뉴 수요 예측은 비용 절감과 고객 만족도 향상에 있어 핵심적인 요소로 작용합니다. 최근에는 AI 기술을 활용한 수요 예측이 식음 서비스 운영의 새로운 해법으로 주목받고 있으며, 정형화된 과거 매출 데이터와 외부 요인을 함께 분석하는 방식이 빠르게 확산되고 있습니다.

이번 해커톤은 리조트 내 식음업장에서의 실전 수요 예측 문제를 AI로 해결해보는 것을 목표로 합니다. Aimers 여러분들은 실제 식음업장에서 수집된 판매 데이터를 기반으로, 각 메뉴가 1주일 동안 얼마나 판매될지를 예측하는 모델을 개발하게 됩니다. 이를 통해 데이터 기반 의사결정이 리조트 운영에 어떤 가치를 더할 수 있는지를 직접 체감할 수 있을 것입니다.

[주제]

리조트 내 식음업장 메뉴별 1주일 수요 예측 AI 모델 개발

[설명]

리조트 식음업장에서 수집된 과거의 메뉴별 판매 데이터를 기반으로, 향후 1주일간 각 메뉴의 예상 판매량을 예측하는 AI 모델을 개발

※ 온라인 해커톤(Phase2)에서는 식음업장의 메뉴 별 매출 규모 데이터만 제공됩니다.

온라인 해커톤(Phase2)에서 교육생들의 문제 해결 능력을 검증하여 오프라인 해커톤(Phase3)에 진출자(약 100명)를 선발하기 위한 과정입니다.

오프라인 해커톤(Phase3)은 1박 2일간 오프라인으로 진행되며, 온라인 해커톤(Phase2)과 주제는 동일합니다.

[주최 / 주관]

- 주최 : LG AI 연구원

- 주관 : 데이콘
- 참여 : 한경닷컴

[참가 자격]

LG Aimers 7기 교육생이라면 누구나 참가 가능

※ Phase 1(온라인 AI 교육) 이수와 관계없이 참가 가능

※ 참여 버튼을 누른 뒤 팝업 창에서 성명, LG Aimers 가입 메일을 기재 (데이콘 계정 메일과 LG Aimers 메일이 일치할 필요 없음)

※ Phase 3(오프라인 해커톤)는 Phase1과 Phase2를 모두 이수한 교육생만 참가 가능합니다. Phase1을 이수하지 않은 경우, Phase2에서 우수한 성적을 거두었더라도 Phase3에 참여할 수 없습니다.

※ 참여 관련 문제가 있는 LG Aimers 교육생의 경우 dacon@dacon.io로 문의

규칙

1. 리더 보드

- **평가 산식** : 식음업장 별 가중치가 있는 SMAPE

$$\text{Score} = \sum_s w_s \cdot \left(\frac{1}{|I_s|} \sum_{i \in I_s} \left(\frac{1}{T_i} \sum_{t=1}^{T_i} \frac{2|A_{t,i} - P_{t,i}|}{|A_{t,i}| + |P_{t,i}|} \right) \right)$$

- s : 식음업장명
- w_s : 식음업장 s 의 가중치 (비공개)
- I_s : 식음업장 s 에 속한 품목 컬럼 집합
- T_i : 품목 i 에서 유효한 날짜 수 ($A_{t,i} \neq 0$)
- $A_{t,i}, P_{t,i}$: 날짜 t , 품목 i 의 실제값과 예측값

◆ 각 식음업장 별 가중치가 존재하며, '담하'와 '미라시아'는 다른 업장보다 높은 가중치로 반영됩니다. (단, 업장 별 가중치 값은 공개하지 않습니다.)

◆ 실제 매출 수량이 0인 경우에는 평가 산식 계산에 반영되지 않습니다.

- **Public score** : 전체 테스트 데이터 샘플 중 사전 샘플링된 50%

- Private score : 전체 테스트 데이터 샘플 100%

2. 평가 방식

- LG Aimers 수료 조건
- Phase1을 이수하고 Phase2의 Public Score (0.711046) 또는 Private Score (0.693935) 이하
- 기준 점수는 Baseline 모델 성능을 기준으로 선정
- **1차 평가** : 리더보드 Private Score 100%
- Private Score는 대회 종료 후 일괄 채점하여 공개
- 동점자의 경우, 기존 리더보드 순위 산정 방식을 따름 [\[링크\]](#)의 '리더보드 점수' 부분을 참고
- **2차 평가** : 오프라인 해커톤(Phase3) 진출을 희망하는 팀은 코드 제출 후 코드 검증
- Private 상위팀(약 100명)은 코드 및 PPT 필수 제출 대상
- 코드 제출과 검증을 모두 통과한 Private 상위팀(약 100명)이 오프라인 해커톤 (Phase3) 진출

3. 개인 또는 팀 참여 규칙

- 개인 또는 팀을 이루어 참여할 수 있습니다.
- 개인 참가 방법: 팀 신청 없이, 자유롭게 제출 탭에서 제출 가능
- 팀 참가 방법: 팀 탭에서 가능, 상세 내용은 팀 탭에서 팀 병합 정책 확인
- 팀 구성 방법: 팀 페이지에서 팀 구성 안내 확인
- 팀 최대 인원: 5명
- 동일인이 개인 또는 복수 팀에 중복하여 등록 불가

4. 외부 데이터 및 사전 학습 모델 관련 규칙

1) 사전학습모델 사용 가능 범위

- 공식적으로 가중치가 공개되었으며, 최소한 비상업적 이용이 허용된 오픈소스 라이선스 (MIT, Apache 2.0 등) 하에 배포된 모델만 사용 가능합니다.
- 해당 조건을 만족하지 않는 모델 및 가중치는 사용할 수 없습니다.

2) API 사용 제한

- 원격 서버 기반의 API 형태로만 접근 가능한 모델(OpenAI API, Gemini API 등)은 사용이 불가합니다.

- 모든 모델은 로컬 환경에서 직접 실행 가능해야 하며, 외부 서버에 의존하는 방식은 제한됩니다.

3) 외부 데이터 사용 금지

- 경진대회에서 제공하는 공식 데이터 외의 외부 데이터는 사용할 수 없습니다.

5. 시계열 예측 관련 Data Leakage 방지 규칙 [\[링크\]](#)

1) 평가 데이터는 학습에 사용할 수 없습니다.

- 평가 Input(28일), Target(7일)은 어떤 경우에도 모델 학습에 활용할 수 없습니다.
- Pseudo Labeling 등 추론 결과를 이용한 재학습도 불가합니다.
- 평가 데이터는 오직 추론 시점에서 입력으로만 사용할 수 있습니다.

2) 추론 시 평가 Input으로 제공된 28일 외의 데이터를 추가로 사용할 수 없습니다.

- 각 평가 샘플에는 Input으로 28일간의 시계열 데이터만 제공되며, 예측 시에는 해당 구간만을 모델 입력으로 사용해야 합니다.
- 평가 시점에서 Lookback 기간을 임의로 확장하거나, 추가적인 과거 데이터를 연결하여 사용하는 것은 허용되지 않습니다.
- 즉, 모든 평가 샘플은 제공된 28일 데이터를 기준으로만 예측이 이루어져야 하며, 모델 구조나 데이터 처리 방식에 관계없이 28일을 초과한 입력 사용은 금지됩니다.

3) 평가 샘플은 서로 독립적으로 추론해야 합니다.

- 하나의 평가 샘플 결과나 입력을 다른 샘플의 예측에 사용하는 것은 금지됩니다.
- 모든 샘플은 각자의 Input(28일)만을 사용해 개별적으로 예측되어야 합니다.

4) 추론 시점 이후 정보는 사용할 수 없습니다.

- 각 평가 샘플의 추론 시점은 Input의 마지막 날짜입니다.
- 이 시점을 기준으로 이후의 데이터(예측 대상 포함)는 모두 미래로 간주되며 활용할 수 없습니다.

5) 외부 데이터는 사용할 수 없지만, 도메인 지식은 활용할 수 있습니다.

- 대회에서 제공한 데이터 외의 외부 데이터, 크롤링, API 호출 등은 금지됩니다.
- 다만, 도메인 지식 기반의 정보(예: 공휴일, 요일 등)는 활용 가능합니다. 다만, 도메인 지식 기반의 정보를 습득할 수 있는 시점도 추론 시점에 유의하여 활용할 수 있어야 합니다.
- 예: "5월 5일은 공휴일이다", "일요일은 주말이다" → 추론 시점 이전에 알 수 있는 도메인 지식 기반의 정보이므로 사용 가능

6. 코드 및 PPT 제출 규칙

- 온라인 해커톤(Phase2) 종료 후 오프라인 해커톤(Phase3) 진출 희망 팀은 아래의 양식에 맞추어 코드를 dacon@dacon.io

로 기한 내에 제출

- 코드에 '/data' 데이터 입/출력 경로 포함
- 코드 파일 확장자: .R, .rmd, .py, .ipynb
- 코드와 주석 인코딩: UTF-8
- 모든 코드는 오류 없이 실행되어야 함(라이브러리 로딩 코드 포함)
- 전체 프로세스를 가독성 있게 정리하고 주석을 포함하여 하나의 파일로 제출
- 개발 환경(OS) 및 라이브러리 버전 기재
- 추론에 사용된 모델 파일 (pth 등) 다운로드 링크 혹은 메일 첨부
- 사용한 사전 학습 모델의 논문 링크, 출처를 기재
- 자유 형식의 솔루션 PPT 필수 제출
- 팀원들의 오프라인 해커톤(Phase3) 참가 여부 필수 기재

7. 유의 사항

- 1일 최대 제출 횟수: **5회**
- 사용 가능 언어: Python
- 모든 학습, 추론의 과정 그리고 추론의 결과물들은 정상적인 코드를 바탕으로 이루어져야 하며, 비정상적인 방법으로 얻은 제출물들은 적발 시 규칙 위반에 해당됩니다.
- 대회 규칙 위반 사항이 적발되는 경우, 최대 실격에 해당할 수 있습니다.
- 최종 순위는 선택된 파일 중에서 채점되므로 참가자는 제출 창에서 자신이 최종적으로 채점 받고 싶은 파일을 **1개**를 선택해야 함
- 대회 직후 공개되는 Private 랭킹은 예선 최종 순위가 아니며 코드 검증 후 본선 진출자가 결정됨
- 데이콘은 부정 제출 행위를 금지하고 있으며 데이콘 대회 부정 제출 이력이 있는 경우 평가가 제한됩니다. 자세한 사항은 아래의 링크를 참고해 주시기 바랍니다.

<https://dacon.io/notice/notice/13>

8. 토론(질문)

- 대회 운영 및 데이터 이상에 관련된 질문 외에는 답변을 드리지 않고 있습니다. 기타 질문은 토론 페이지를 통해 자유롭게 토론해 주시기 바랍니다.

- 데이콘 답변을 희망하는 경우 토크 게시글 댓글로 질문을 올려 주시기 바랍니다.

예) [DACON 답변 요청] 시상식은 언제 열리나요?

일정

Phase2. 온라인 해커톤 일정

- 대회 기간 : 2025년 08월 01일(금) 10:00 ~ 2025년 08월 25일(월) 10:00
- 팀 병합 마감 : 2025년 08월 18일(월) 23:59
- 대회 종료 : 2025년 08월 25일(월) 10:00
- 코드 및 PPT 제출 : 2025년 08월 25일(월) 12:00 ~ 2025년 08월 27일(수) 10:00
- 코드 검증 : 2025년 08월 27일(수) 10:00 ~ 2025년 09월 01일(월) 14:00
- 오프라인 해커톤(Phase3) 진출자 발표 : 2025년 09월 01일(월) 14:00

※ 오프라인 해커톤(Phase3)은 1박 2일 오프라인으로 별도 대회 페이지에서 진행됩니다.
(9/6-9/7 예정)

※ 세부 일정은 대회 운영상황에 따라 변동될 수 있습니다.

상금

[🏆 총 상금 1000만 원]

🥇 1위 - 500만 원 (고용노동부장관상)

🥈 2위 - 300만 원 (LG AI연구원장상)

🥉 3위 - 200만 원 (한경닷컴 사장상)

- 본 온라인 해커톤(Phase2)은 오프라인 해커톤(Phase3) 진출 팀을 선정하기 위한 과정으로 별도의 상금은 없으며, 오프라인 해커톤(Phase3) 상위 3팀에게 상금이 수여됩니다.
- 온라인 해커톤(Phase2)을 통해 코드 제출과 코드 검증을 통과한 Private 상위팀(약 100명)이 오프라인 해커톤(Phase3) 진출 팀으로 선정됩니다.

※ 제세공과금은 개인 부담

데이터

Dataset Info.

- **train [폴더] :**
 - **train.csv [파일]**
 - 2023.01.01 ~ 2024.06.15의 영업장명_메뉴명별 매출 수량 정보
 - 영업일자
 - 영업장명_메뉴명
 - 매출수량

| 영업일자 | 영업장명_메뉴명 | 매출수량 |
|------------|--------------------|------|
| 2023-01-01 | 느티나무 셀프BBQ_1인 수저세트 | 0 |
| 2023-01-02 | 느티나무 셀프BBQ_1인 수저세트 | 0 |
| 2023-01-03 | 느티나무 셀프BBQ_1인 수저세트 | 0 |
| 2023-01-04 | 느티나무 셀프BBQ_1인 수저세트 | 0 |
| 2023-01-05 | 느티나무 셀프BBQ_1인 수저세트 | 0 |
| 2023-01-06 | 느티나무 셀프BBQ_1인 수저세트 | 0 |
| 2023-01-07 | 느티나무 셀프BBQ_1인 수저세트 | 0 |
| 2023-01-08 | 느티나무 셀프BBQ_1인 수저세트 | 0 |
| 2023-01-09 | 느티나무 셀프BBQ_1인 수저세트 | 0 |
| 2023-01-10 | 느티나무 셀프BBQ_1인 수저세트 | 0 |

- **test [폴더] :**

- **TEST_00.csv ~ TEST_09.csv [파일]**

- 2025년의 특정 시점(28일)의 영업장명_메뉴명별 매출 수량 정보
- 영업일자
- 영업장명_메뉴명
- 매출수량

| 영업일자 | 영업장명_메뉴명 | 매출수량 |
|------------|--------------------|------|
| 2024-06-16 | 느티나무 셀프BBQ_1인 수저세트 | 2 |
| 2024-06-17 | 느티나무 셀프BBQ_1인 수저세트 | 0 |
| 2024-06-18 | 느티나무 셀프BBQ_1인 수저세트 | 0 |
| 2024-06-19 | 느티나무 셀프BBQ_1인 수저세트 | 0 |
| 2024-06-20 | 느티나무 셀프BBQ_1인 수저세트 | 4 |
| 2024-06-21 | 느티나무 셀프BBQ_1인 수저세트 | 8 |
| 2024-06-22 | 느티나무 셀프BBQ_1인 수저세트 | 39 |
| 2024-06-23 | 느티나무 셀프BBQ_1인 수저세트 | 7 |
| 2024-06-24 | 느티나무 셀프BBQ_1인 수저세트 | 0 |
| 2024-06-25 | 느티나무 셀프BBQ_1인 수저세트 | 0 |

- **sample_submission.csv [파일] - 제출 양식**

- 각 영업장명_메뉴명의 TEST 파일별 +1일, +2일,..., +7일의 매출수량 예측 결과
- 영업일자 : TEST_00+1일, TEST_00+2일, TEST_00+3일 ... TEST_09+1일, TEST_09+2일, TEST_09+7일