OO 대회 인공지능 해커톤
모델 설명서

팀 명 : **O O O**
2023년 **8**월 **OO**일

1. 라이브러리 및 데이터 **(Library & Data)**

-
-

(예시)
- 데이터를 불러올 때 인코딩 문제 때문에 제공 파일을 텍스트 파일로 변환 후 불러옴
- 데이터 테이블 형태인 *fread*로 텍스트 파일을 읽음

code

작성요령

- 라이브러리 및 데이터에 대한 분석 결과에 대한 내용 서술

2. 데이터 전처리 **(Data Cleansing & Pre-Processing)**

-
-

(예시)
- 날짜와 시간을 합친 *'transacted_datetime'* 변수를 생성
- 환불은 로그를 씌웠을 때 무한대가 나오기 때문에 제거하기로 결정

code

작성요령

- 데이터의 정제/통합/정리/변환 등 데이터 전처리 결과에 대한 내용 서술

3. 탐색적 자료 분석 **(Exploratory Data Analysis)**

-
-

(예시)
- 상점별 시계열 데이터의 *AR*값과 *adp.test p-value*값을 계산
- *ar*차수와 *adf_p*의 *boxplot*을 보면 *ar*차수는 대부분의 데이터가 *2*이하이고,대부분의 데이터가 차분이 필요해 보임

code

작성요령

- 다양한 각도에서 관찰하고 이해하는 과정으로 데이터를 분석하기 전에 그래프나 통계적인 방법으로 자료를 직관적으로 바라본 결과에 대한 내용 서술

4. 변수 선택 및 모델 구축 **(Feature Engineering & Initial Modeling)**

-
-

(예시)
- 매출액 변동계수 변수 생성
- 최종 *feature*로 *store_id,amount,yyyymm(*년월*), CV(*변동계수*)*선택

code

작성요령

- 기법 선택 및 변수 선정에 대한 내용 서술
- 구축된 딥러닝 모델에 대한 내용 서술

5. 모델 학습 및 검증 **(Model Tuning & Evaluation)**

-
-

(예시)
- *auto.arima,stl,ets*모형을 각각 *0.33*의 동일한 가중치로 적합
- *auto.arima* 파라미터만 *ndiffs*함수와 사전 *ar*값 조사를 통해 *max.p*를 *2*로 고정

code

작성요령

- 제공된 데이터셋 기반 모델 학습 및 검증에 대한 내용 서술

6. 결과 및 결언 **(Conclusion & Discussion)**

-
-

(예시)
- 로그를 취했을 때 과대계상되는 상점들을 명확하게 구분방법을 알 필요가 있음
- 한달 단위로 리샘플링 했지만 다른 방식으로도 해볼 필요가 있음

code

작성요령

- 내용 중에서 가장 중요한 점을 다시 한 번 강조하고
- 경진대회 이후 모델 보완 및 추가 연구에 대한 방향성 등 서술