

주차예측프로젝트 보고

T1F4

구지윤, 송수인, 조성진, 조형진, 최민재

목차

1. 문제 정의
2. 오산시 공영주차장 특징
3. 사용데이터
4. EDA
5. 데이터 전처리
6. 모델링
7. 기대효과 및 향후 연구

문제 정의

오산시의 문제점

- 현재 오산시는 도로증가량에 비해 차량증가율의 비율이 커 주차장 확보에 어려움을 겪고있음
- 전통시장의 경우 좁고 일방통행이라는 단점이 있어 이용객이 불편함을 겪고있음
- 주차장 증설엔 경제적인 부담이 크기 때문에 증설보다는 혼잡도를 예측하여 주차장의 수요를 분산시키고자 함

지정한 날짜(오늘)를 포함한 일주일간 각 주차장의 현재 주차대수를 예측하는 모델

간격은 15분이며, 현재 시간을 기점으로 1시간 뒤까지의 주차장 혼잡도를 표정 아이콘으로 시각화하여 나타내기로 함.

*왜 15분인가? : 오산시는 자가용 이용 시 공영주차장까지 최대 15분 소요, 또한 15분 단위인 타 공공데이터와의 호응성과 30분과 한시간 단위로 환산이 용이하기 때문

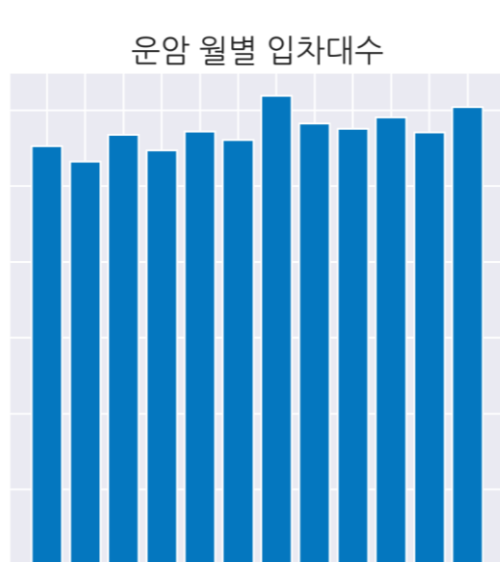
오산시 주차장 특징

데이터 EDA를 통해 주차장의 특징을 확인

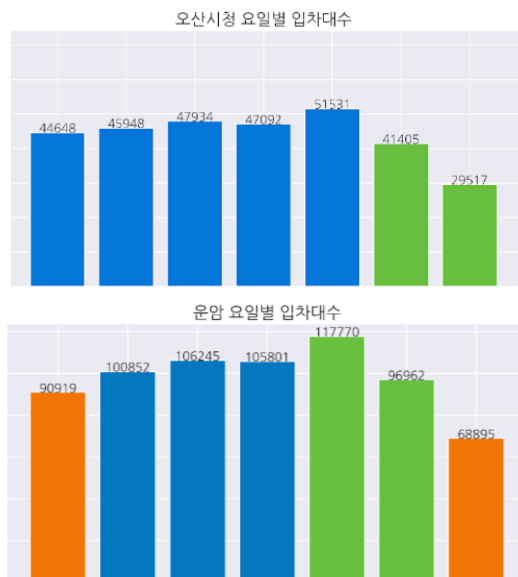
이름		주요 특징	참고사항
주차장1	오산시청부설주차장	관공서(시청) 인근	아침에는 주차시간이 7시간 이상으로 출근차량들로 구성, 주간에는 주차시간이 1시간으로 민원인 차량으로 예상. 이면주차 가능 저녁 시간에 입출차가 많다. 출입구가 일직선으로 연결되어 있어 입출차 시간의 간격이 짧은 통과차량이 있는 것이 특징.
주차장2	운암공영주차장	먹자골목 인근	
주차장3	오색시장1공영주차장	전통시장 인근	
주차장4	오색시장2공영주차장	전통시장/주택가 인근	

EDA

- 연별 입출차 EDA: 데이터의 기간이 1년이라 고려하지 않음
- 월별 입출차 EDA: 월 별로 일수가 다른 것을 고려해도, 모든 주차장의 월 별 입/출차는 비슷함 (그림1 참고)
- 요일별 입출차 EDA: 주차장1 (관공서)는 주말이용객이 적고, 주차장3,4(전통시장)는 주말이용객이 많고, 주차장2(먹자골목)는 타 주차장과 다르게 금/토 이용객이 많음 (그림2 참고)
- 일별 입출차 EDA: 주차장3,4(전통시장)이 오산시 전통시장의 오일장(3일, 8일의 장터)때 이용객이 많음 (그림3 참고)



[그림1]



[그림2]

대각선: 주차장1
위쪽: 주차장3,4
왼쪽: 주차장2

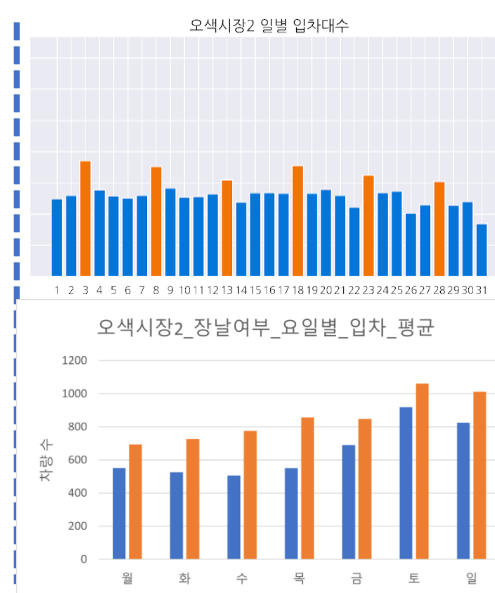
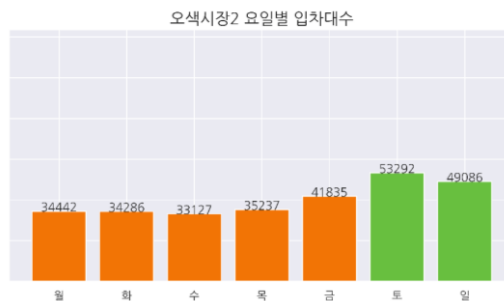


그림3:
주황색: 오일장

사용데이터

1) 오산시 측에서 제공받은 데이터

데이터 명	기간	대상	설명
오산요청_0821.csv	2022-07-01 ~ 2023-06-30 (입차기준)	주차장 4개	2022년도 하반기 ~ 2023년 상반기
	2022-07-01 ~ 2023-07-31 (출차기준)		. 주차장 명, 차량번호(비식별) . 입차시각, 출차시각

2) 사용한 추가 공공데이터: 기상자료개방포털>종관기상예측(ASOS) 및 공공기관 기상관측 데이터

지점	일시	기온	강수량
550	2022-07-01 01:00	23	0
550	2022-07-01 02:00	23	0
550	2022-07-01 03:00	22.6	0

일부 결측치: 근접 지역인 수원, 서탄면(1405), 남촌(446)의 데이터를 활용하여 보완

데이터 전처리: 결측치, 장기주차, 명절

모든 주차장의 결측치 비율은 1%대로, 주로 22년12월13일~22년12월15일에 많은 결측치가 관측됨.

비식별화된 차량번호 중 특정차량이 하루 평균 4번 이상, 다양한 주차장에 입차기록이 있으며, 출차기록은 모두 결측치로 확인됨.

→ 주차장을 단기적으로 이용하거나, 통과하기 위해 사용했을 것으로 사료되어 분석에서 해당 차량번호 행은 제외함. (총 5957개)

다음으로, 결측치 개수를 내림차순으로 그리면 아래와 같이 결측치가 관측된다.

주차장명	일시	결측치 수
주차장4(전통시장2)	12월14일 12시~14시	164개
주차장2(먹자골목)	6월1일~8일 8시	87개
주차장2(먹자골목)	9월 9일 16시	20개
주차장3(전통시장1)	9월 9일 17시	48개

주차장4의 해당일시는 결측치가 50% 이상을 차지하여 데이터에서 제외하였다.

주차장2는 통과주차장의 특징이 있어 출퇴근 통과차량이 많아 데이터에서 제외 (주차대수 예측에 영향력 없음)

9월9일은 추석이므로 제외

데이터 전처리: 결측치, 장기주차, 명절

이외의 결측치는 일별 평균 비율이 0.5% 이하로 발생되어 결측치를 제외하는 대신에, 과대추정을 하는 방식으로 진행하기로 함.

단, 아래 테이블과 같이 데이터에서의 마지막날에서의 출차결측은, 아직 출차하지 않은 차량일 수 있다.

-> 먹자골목 주차장은 먹자골목 인근이기 때문에 늦은시간에 출차하는 차량이 많음

-> 관공서 주차장은 관공서 인근이지만, 저녁시간대부터는 자유롭게 이용하기 때문에 늦은시간 출차 가능

그러므로, 데이터를 불러올 때 마지막날 결측치는 제외하지 않도록 함수화하였다.

입차	출차	주차장명
2023-06-30 19:02:03	NaT	주차장2
2023-06-30 19:00:12	NaT	주차장2
2023-06-30 20:56:00	NaT	주차장1

데이터 전처리: 결측치, 장기주차, 명절

주차장 명	주차장1 (관공서)	주차장2 (먹자골목)	주차장3 (전통시장1)	주차장4 (전통시장2)
장기주차(개수)	9,141	2,642	2,411	705
장기주차(비율)	1.48%	0.38%	0.27%	0.92%

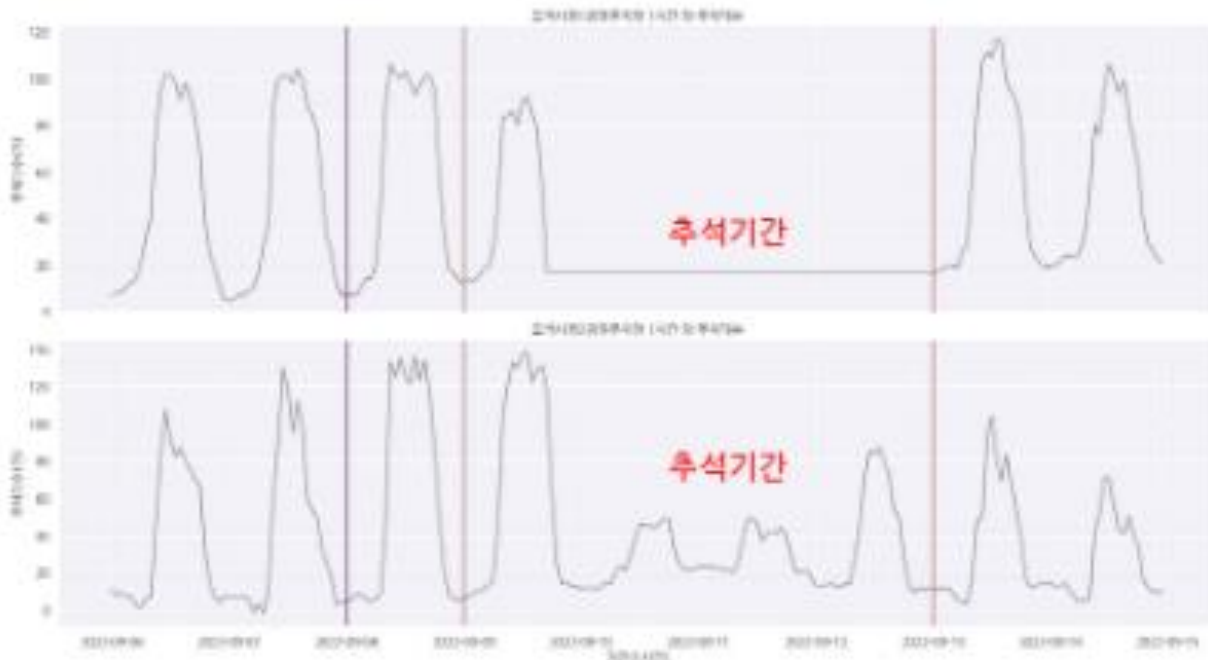
[1] 실무자와 논의 사항

전통시장 1,2 : 모두 이상치
관공서: 관용차량의 가능성 (2023년 9월 기준 35대)

[2] 데이터 상의 한계점

데이터 제공기간 이전 (2022년 7월 1일 이전) 입차했지만
7월 1일 이후에 출차한 장기주차 차량들에 대한 데이터는 포함되지 않음
따라서, 장기주차를 고려하기 어려움

데이터 전처리: 결측치, 장기주차, 명절

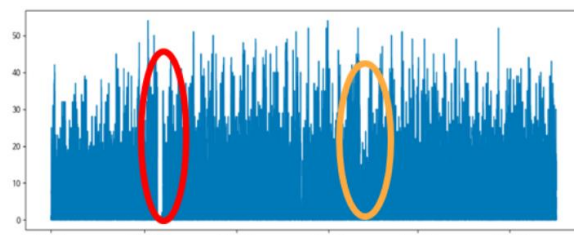
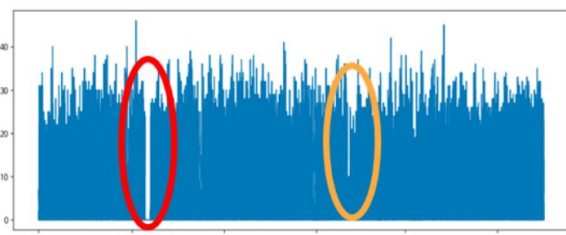
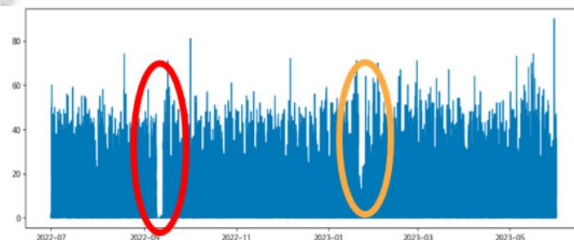
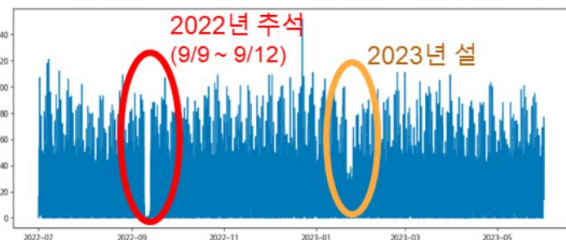


제거대상

1. 명절 직전 비정상적으로 증가한 사용인원
2. 데이터가 제대로 갖춰지지 않은 명절

이유

1. 입출차 데이터의 부재
차단기 관리방식이 확립되지 않아 부정확한 집계
2. 특수성에 비해 적은 빈도 수
명절 특수성을 가진 기간이 전체 데이터 내에 단 두 번
→ 표본이 너무 부족



3. 추가 공영주차장 운영

명절 연휴에만 한시적으로 추가 공영주차장을 운영하였고, 해당 부분이 평소와 다르게 영향을 주었을 것임
(담당부서가 달라 데이터도 없음)

=> (명절+앞 이틀)을 모두 “명절기간”으로 명명하고 학습데이터에서 제외

데이터 전처리: 누적 값 환산

CarNumber	In	Out
2744A3D3HGzaqjHp	2022-07-01 00:14:47	2022-07-01 00:26:47
7B8429D8DddulCDT	2022-07-01 00:21:08	2022-07-01 00:24:27

입출차 분리
시간 순 정렬

입차: +1
출차: -1

15분 단위 재조정

Index	증감	누적
2022-07-01 18:15:00	12	273
2022-07-01 18:30:00	-7	266
2022-07-01 18:45:00	3	269

초깃값

시작시점 남은 주차면수:

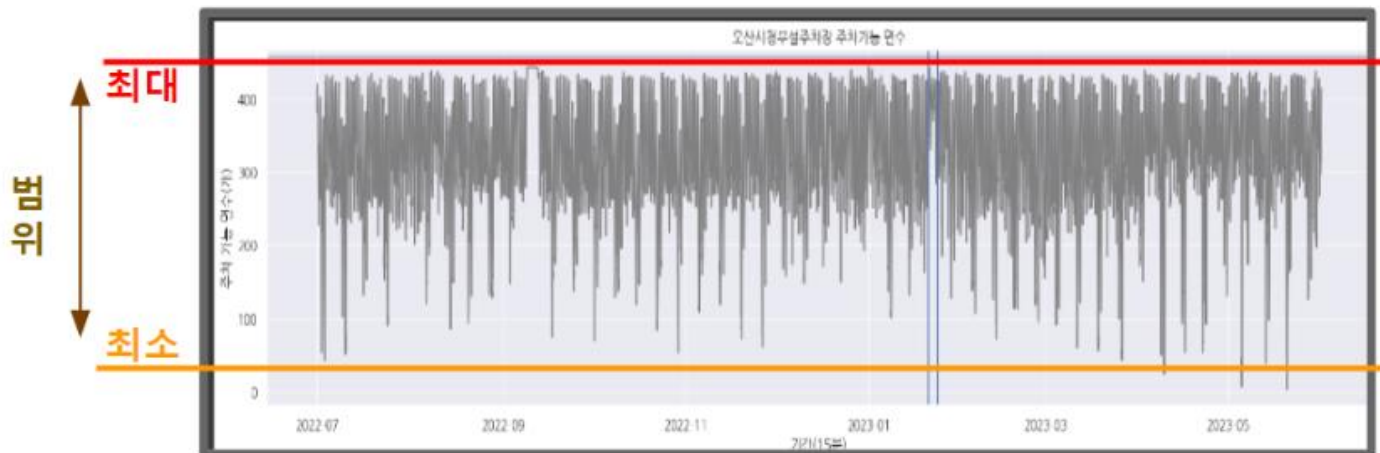
→ 보수적으로 가능한 크게!

제약 1) 범위 \leq 총 주차면

제약 2) 최소 주차대수 > 0

제약 3) 시작시점 남은 주차면수
≒ 실측값 (or 실무자 직관)

[남은 주차면수 개형 예시]



관공서주차장	먹자골목주차장	전통시장1주차장	전통시장2주차장
150대	50대	10대	20대

변수설정 (X변수)

다음은 데이터 분석과 모델의 성능향상을 위해 오산시 제공 주차데이터로부터 생성한 변수다.

변수 명	의미	참고사항
timing cos	하루 중 해당하는 시간대를 cyclic하게 표현	
timing sin	하루 중 해당하는 시간대를 cyclic하게 표현	
요일	날짜에서 '요일' 추출	
주말	주말 여부를 나타낸 변수	
장날	전통시장의 오일장 장날 여부를 나타낸 변수	매월 3, 8일 오일장 (전통시장 주차장에만 활용)
공휴일	공휴일 여부를 나타낸 변수	
3일 이상	3일이상 연휴 여부를 나타낸 변수	주말과 공휴일이 겹친 경우 Ex) 2022-08-15(월)
먹자골목 요일	일월 / 화수목 / 금토	묵음으로 기간 표현 (먹자골목 주차장 예측에만 활용)
7일전 주차면수	7일전의 주차면수를 나타낸 변수	주차장1, 주차장3에 사용
주차시간 통계량	차량에서 주차한 시간으로 (출차 - 입차 시간)표현	장기주차의 기준은 24시간 이상으로 생각.

변수설정 (X변수)

제공 외의 추가로 모은 공공데이터로는 강수량과 기온데이터를 사용했다.

변수 명	의미	참고사항
온도	15분 별 온도	시간당 값들을 위아래 평균화
강수량	15분당 누적 강수량	1시간 단위 누적 강수량을 4로 나눠 15분 단위에 채워 넣었다
30분전_온도	30분전 온도	
30분전_강수량	30분전 강수량	

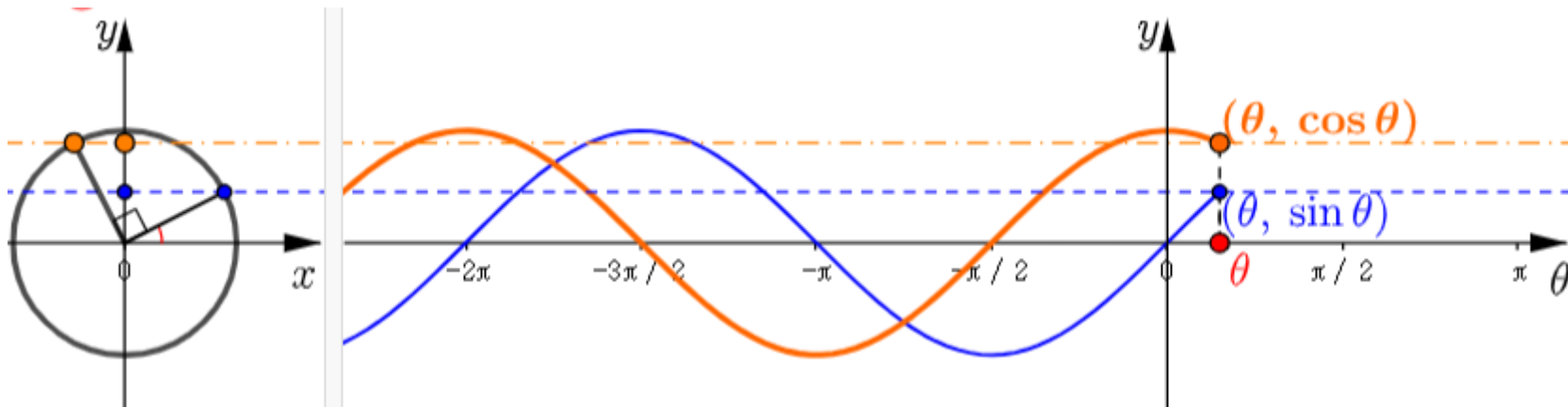
현재 기상정보: 어떤 주차장을 활용할 지 (야외/실내)

30분전 기상정보: 출발시점에서 교통수단 선택에 영향을 줌 (자가용/대중교통)

시간대 변수설정 (X변수)

Timing sin, cos
24시간 X 4 = 96 구간

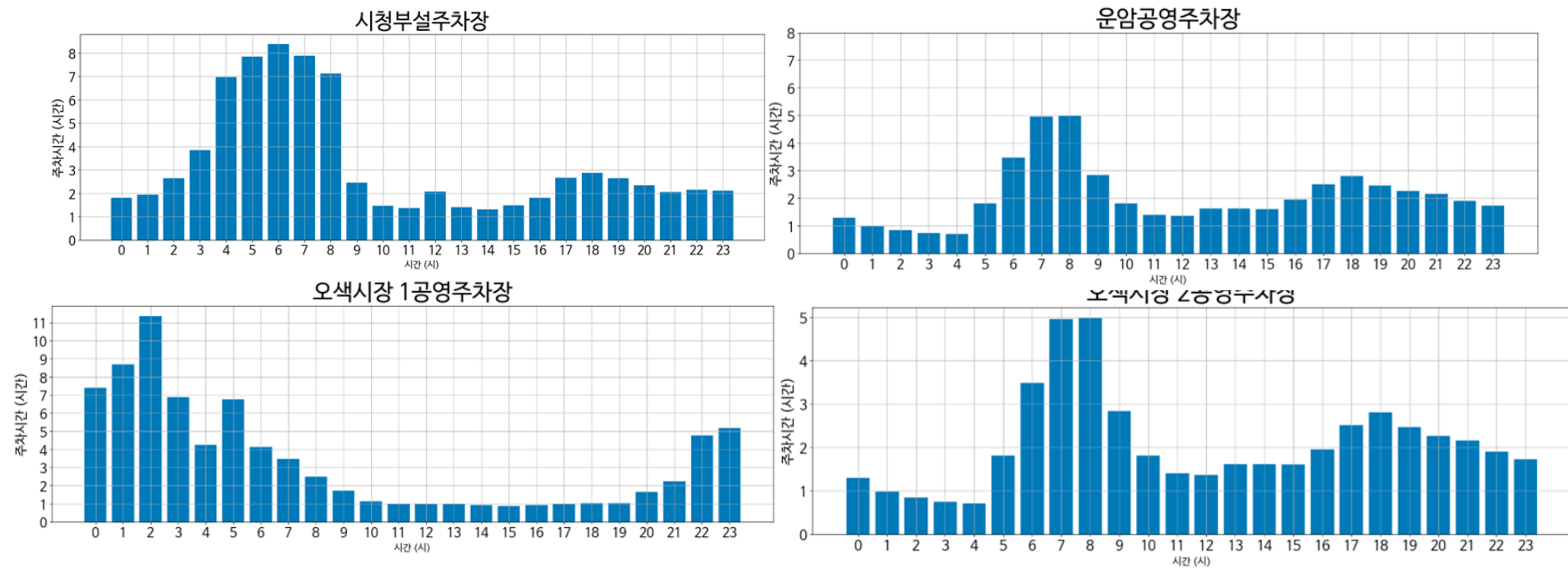
변수명	의미
timing_sin	$\sin(\text{시간대} * 2\pi / 96)$
timing_cos	$\cos(\text{시간대} * 2\pi / 96)$



하루 24시간을 15분 단위(시간별 4구간) 총 96구간으로 나누어 하루 (시간대* $2\pi/96$)의 형태로 주기성 표현

11시 45분 구간과 다음날 00시 00분 구간 사이를 연결하여 cyclic한 속성을 반영 가능

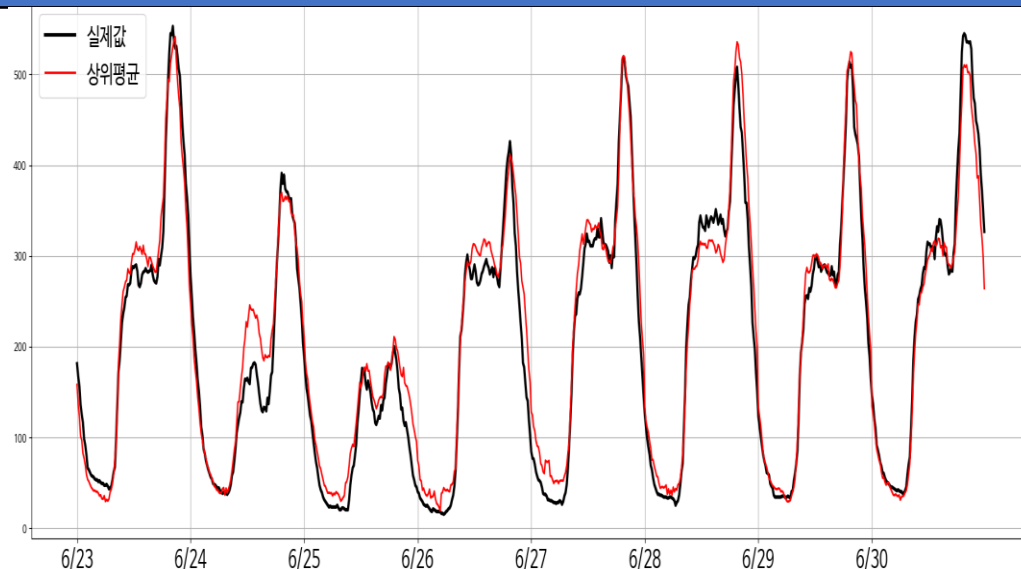
시간대 변수설정 (X변수)



변수명	의미	참고사항
mins_taken_mean	해당 시간대 입차차량 평균 주차시간	사분위수 값들로 너무 극단적인 값들 영향을 배제하며 높은 값과 낮은 값을 표현
mins_taken_std	해당 시간대 입차차량 주차시간 표준편차	
mins_taken_Q1	주차시간대에 제 1사분위수	
mins_taken_Q3	주차시간대에 제 3사분위수	

모델링에서 변수선택을 안하는 이유

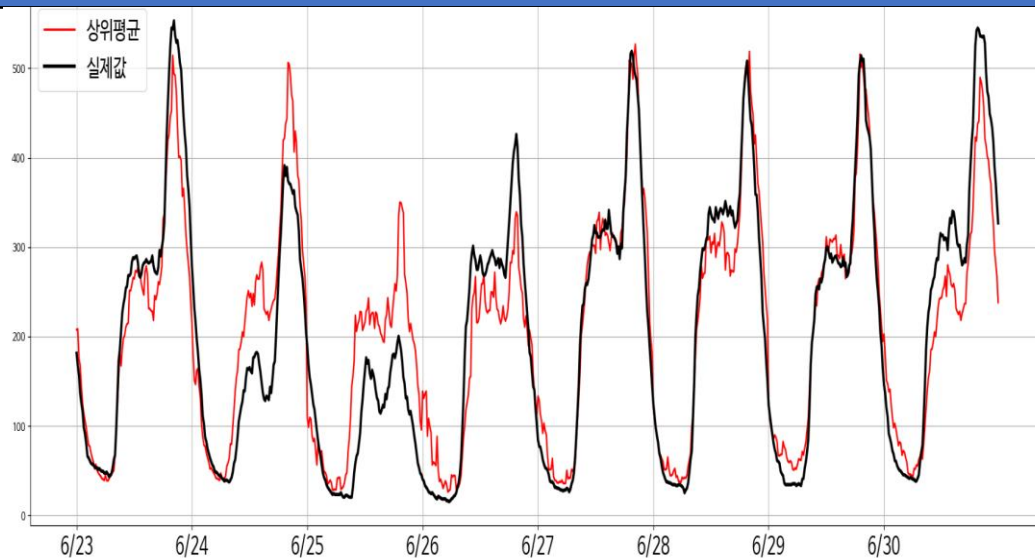
변수선택 X



혼잡한 극단값 근처 예측력 준수

세밀한 개형으로 세밀한 변동 예측

변수선택 O



혼잡한 극단값 근처 예측력 부족

단순한 개형으로 세밀한 변동 예측하지 못함

모델선정

딥러닝

단순한 데이터와 피쳐 데이터가 적음
=> 머신러닝에 비해 현저히 낮은 성능을 가졌다.

GPU 확보가 불분명함
담당자의 관련지식 부재

시계열

월별 영향(Seasonality) 없음
주차면적 제한으로 인해 추세(Trend) 없음

극단값(혼잡도)를 잘 예측하지 못함.

기본회귀

적합이 잘 안됨

다른 모델에 비해 낮은 설명력



부스팅 / 배깅

1. 편중(bias)과 분산(variance)을 줄일 수 있다.
2. 모델의 과적합 또는 과소적합 방지
3. 더 안정적이고 노이즈가 적다.

모델선정 (XGBOOST)

Model	MAE	MAPE
XGBoost	13	0.2245
LightGBM	13.9	0.2505
CatBoost	14.08	0.2463
Random Forest	15.55	0.2562
Gradient boosting	19.66	0.3339

LightGBM: 데이터 개수가 적어 제외

CatBoost: 범주형 자료에 강점이 있어 제외

Gradient Boosting: XGBoost와 속도의 차이가 있으며, XGBoost가 더 성능이 좋음으로 제외

Random Forest: 메모리 소모가 크며 train data를 추가해도 모델 성능 개선이 어렵다

Weighted MSE Loss, MAE Metric

[가중치]

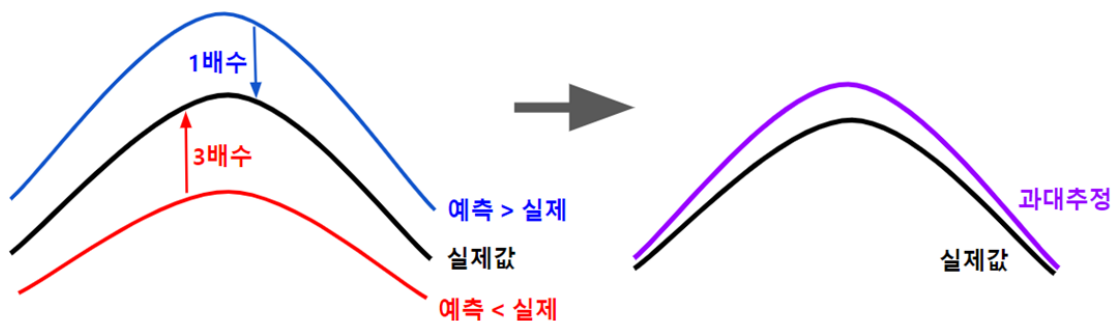
주차장명	관공서	먹자골목	전통시장1	전통시장2
과대 : 과소	1 : 10	1 : 10	1 : 2	1 : 3
주차면수	458	593	105	180

과대추정의 일환으로 예측값이 실제값보다 작은 **과소추정을 최대한 방지**하기 위해 맞춤 손실함수와 평가지표를 만들었다.

(예측값 < 실제값) 의 경우 손실값을 가중 주차장 총 면수를 고려하여 가중치 크기를 결정함

custom loss 뿐만 아니라 평가지표인 custom metric도 정의했다

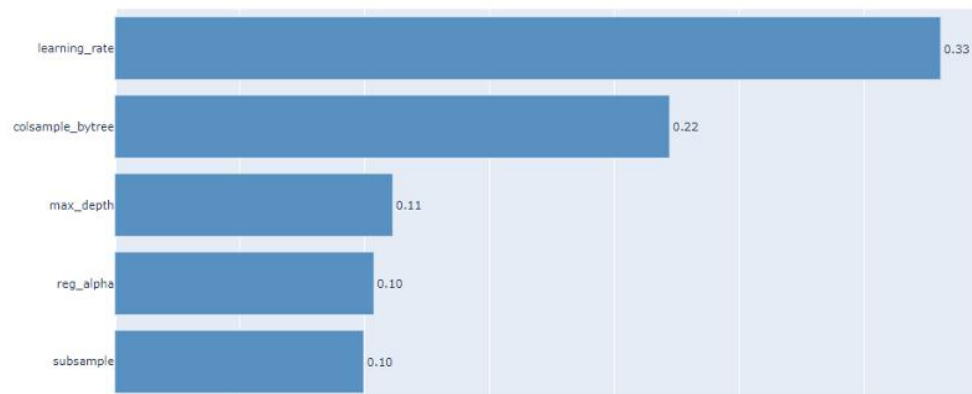
Ex) 시청의 경우 예측값이 10이고, 실제값이 15이면,
손실함수의 loss: 25이 아닌 $25 \times 100 = 2500$
평가지표 metric: 50이 아닌 $5 \times 10 = 500$ 이 된다.



Optuna 하이퍼 파라미터 최적화

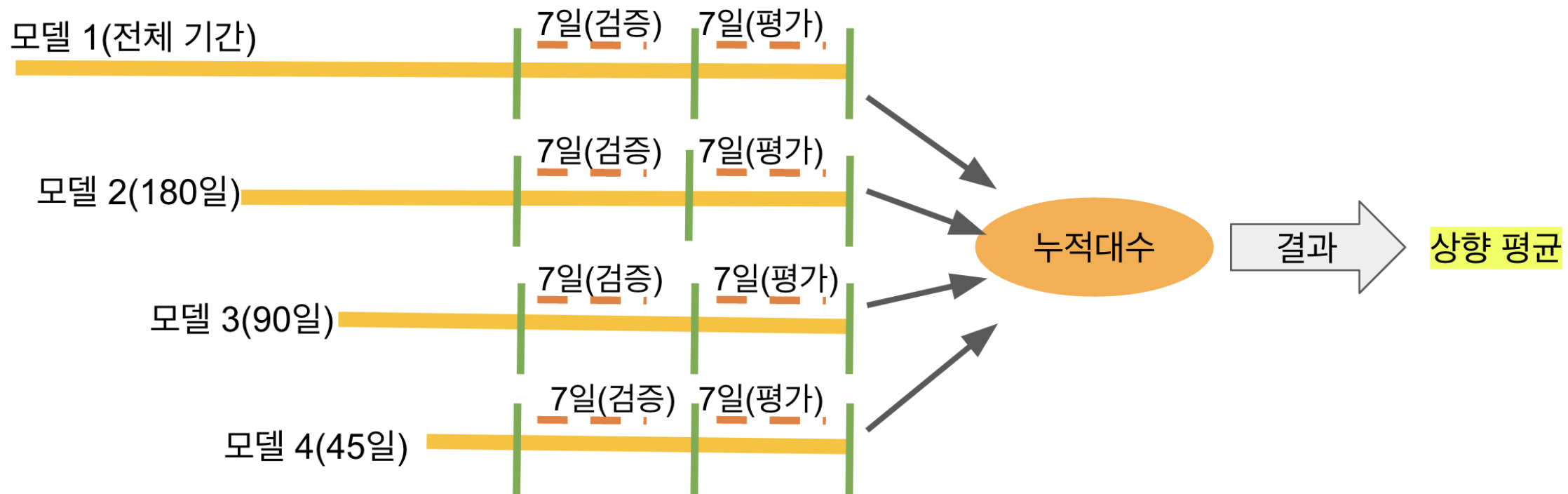
주차장별 특성에 따라 서로 다른 파라미터를 중점적으로 사용

Hyperparameter Importances



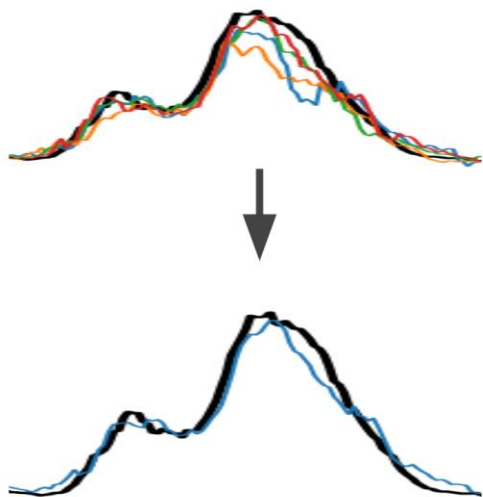
```
[I 2023-09-20 17:21:31,684] Trial 2 finished with value: 52.53271611531576 and parameters: {'max_depth': 14, 'learning_rate': 0.5707021081897535, 'min_child_weight': 3, 'gamma': 0.8069936843273351, 'subsample': 0.37743072594509297, 'colsample_bytree': 0.987, 'reg_lambda': 0.7580258109335488, 'random_state': 5}. Best is trial 2 with value: 52.53271611531576.
[I 2023-09-20 17:21:32,701] Trial 3 finished with value: 24.568097682226274 and parameters: {'max_depth': 8, 'learning_rate': 0.4671081055890881, 'min_child_weight': 5, 'gamma': 0.560874454761309, 'subsample': 0.3815772835100183, 'colsample_bytree': 0.252, 'reg_lambda': 0.10167762279495893, 'random_state': 4}. Best is trial 3 with value: 24.568097682226274.
[I 2023-09-20 17:21:33,605] Trial 4 finished with value: 397.72039804004487 and parameters: {'max_depth': 3, 'learning_rate': 0.17613092251934026, 'min_child_weight': 8, 'gamma': 0.40429599726034493, 'subsample': 0.9133909878119049, 'colsample_bytree': 0.19152, 'reg_lambda': 0.20951376782300504, 'random_state': 46}. Best is trial 3 with value: 24.568097682226274.
[I 2023-09-20 17:21:34,443] Trial 5 finished with value: 46.036592914944606 and parameters: {'max_depth': 3, 'learning_rate': 0.6713953807941625, 'min_child_weight': 8, 'gamma': 0.24556677369407784, 'subsample': 0.5549530185379313, 'colsample_bytree': 0.888, 'reg_lambda': 0.2297219653344417, 'random_state': 11}. Best is trial 3 with value: 24.568097682226274.
[I 2023-09-20 17:21:35,494] Trial 6 finished with value: 54.76423729033697 and parameters: {'max_depth': 13, 'learning_rate': 0.5192318259040091, 'min_child_weight': 6, 'gamma': 0.4808035846720615, 'subsample': 0.8296978431764451, 'colsample_bytree': 0.5, 'reg_lambda': 0.7913536778599337, 'random_state': 22}. Best is trial 3 with value: 24.568097682226274.
[I 2023-09-20 17:21:36,703] Trial 7 finished with value: 50.4205790247236 and parameters: {'max_depth': 14, 'learning_rate': 0.6869804020601785, 'min_child_weight': 6, 'gamma': 0.3933874405785143, 'subsample': 0.7695105938924603, 'colsample_bytree': 0.9, 'reg_lambda': 0.9603563375966335, 'random_state': 40}. Best is trial 3 with value: 24.568097682226274.
[I 2023-09-20 17:21:37,561] Trial 8 finished with value: 48.17793126333328 and parameters: {'max_depth': 5, 'learning_rate': 0.7463635252001162, 'min_child_weight': 7, 'gamma': 0.24855149774521818, 'subsample': 0.351350904479249, 'colsample_bytree': 0.92, 'reg_lambda': 0.10024716350249549, 'random_state': 44}. Best is trial 3 with value: 24.568097682226274.
```

기간별 모델링

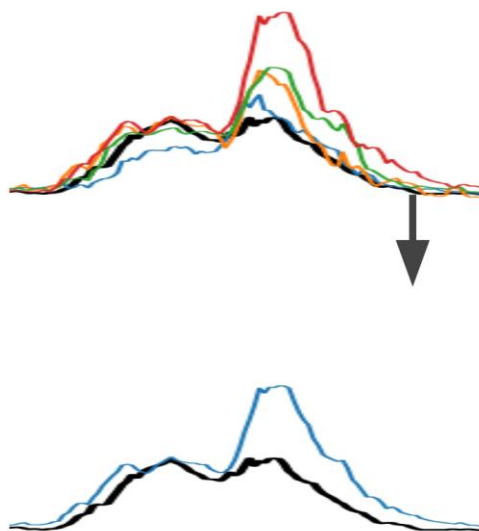


상향평균

[1] 낮은 예측을 방지



[2] 적정수준의 정확도 유지



4개의 모델을 통해 나온 누적대수 중 상위 2개의 값의 평균을 구함.

주차장 이용객의 심리를 반영하여 높게 예측하는 것이 상대적으로 이용자들에게 더 높은 만족감을 줄 수 있을 수 있다고 생각한다.

MLFlow 이용

주차장 4개 X 기간4개 = 모델 16개

MLFlow : 모델의 평가지표와 초매개변수를 서로 확인
할 수 있는 협업툴

DagsHub : MLFlow를 연동하는 원격 레포지토리

장점: 평가지표가 좋은 모델을 받아와 바로 적용 가능
불필요하게 겹치는 작업을 최소화했다.

우측은 Mlflow를 이용하여 훈련한 모델의 아티팩트다.

피처중요도, 선택된 변수, 하이퍼파라미터 등 모델 관련 정보
들을 기록·보관하여 협업환경을 구축



▼ Artifacts

▼ model

- MLModel
- conda.yaml
- model.xgb
- python_env.yaml
- requirements.txt
- feature_importance_weight.jsor
- feature_importance_weight.png

Full Path: mlflow-artifacts/18f7d5bb38724a93b0c... [📄](#) [Register Model](#)

MLflow Model

The code snippets below demonstrate how to make predictions using the logged model. You can also [register it to the model registry](#) to version control.

Model schema

Input and output schema for your model. Predict on a Spark DataFrame: [📄](#)

[Learn more](#)

Name	Type
Inputs (18)	
timing_cos	double

Make Predictions

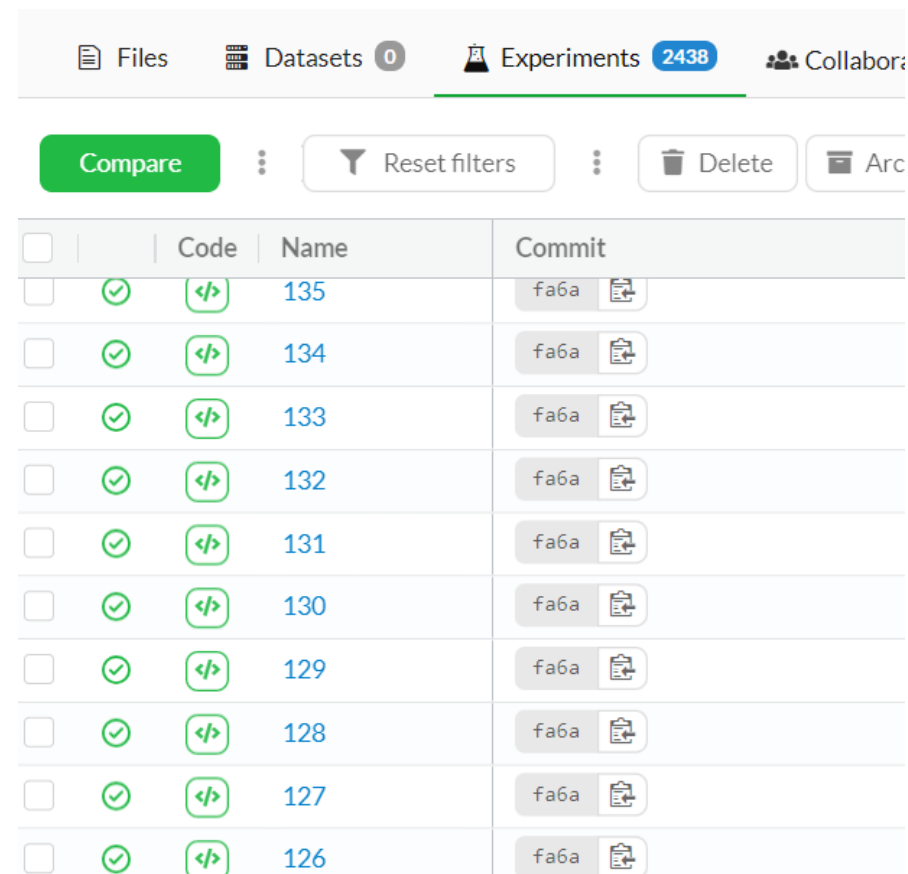
```
import mlflow
from pyspark.sql.functions import struct, col
logged_model = 'runs:/9a0a4cf0939f4b759f2ca113fe27e78c/model'

# Load model as a Spark UDF. Override result_type if the model does not return double values.
loaded_model = mlflow.pyfunc.spark_udf(spark, model_ur
```


MLFlow 이용

```
User@Jiyun MINGW64 ~/Desktop/New folder/source (main)
$ python main.py 2023-06-24
sichung parkinglot is in progress.....
Read Raw Dataset ...
Dataset Imported
Our Prediction Week is from 2023-06-24 to 2023-06-30
Output: X,y : Cols are selected before training/test
[I 2023-09-19 10:03:40,285] A new study created in memory with name: sichung_total_optuna
```

Optuna로 하이퍼파라미터 튜닝 시, 훈련했던 모든 경우의 하이퍼파라미터를 mlflow를 통해 확인할 수 있다 (숫자 0~) 또한 그 중에서 가장 좋은 파라미터를 이용하여 훈련 시, 그 모델은 Train/Val Metric 또한 저장된다.



The screenshot shows the Dagshub Experiments interface. At the top, there are tabs for Files, Datasets (0), Experiments (2438), and Collaborators. Below the tabs, there are buttons for Compare, Reset filters, Delete, and Archive. The main content is a table of experiments.

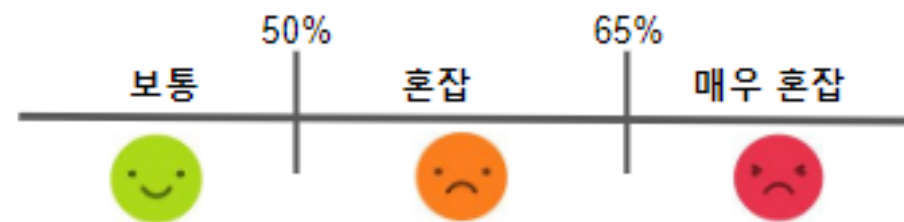
<input type="checkbox"/>		Code	Name	Commit
<input type="checkbox"/>	✓		135	fa6a
<input type="checkbox"/>	✓		134	fa6a
<input type="checkbox"/>	✓		133	fa6a
<input type="checkbox"/>	✓		132	fa6a
<input type="checkbox"/>	✓		131	fa6a
<input type="checkbox"/>	✓		130	fa6a
<input type="checkbox"/>	✓		129	fa6a
<input type="checkbox"/>	✓		128	fa6a
<input type="checkbox"/>	✓		127	fa6a
<input type="checkbox"/>	✓		126	fa6a

[Optuna 튜닝 중의 Dagshub Experiment 화면]

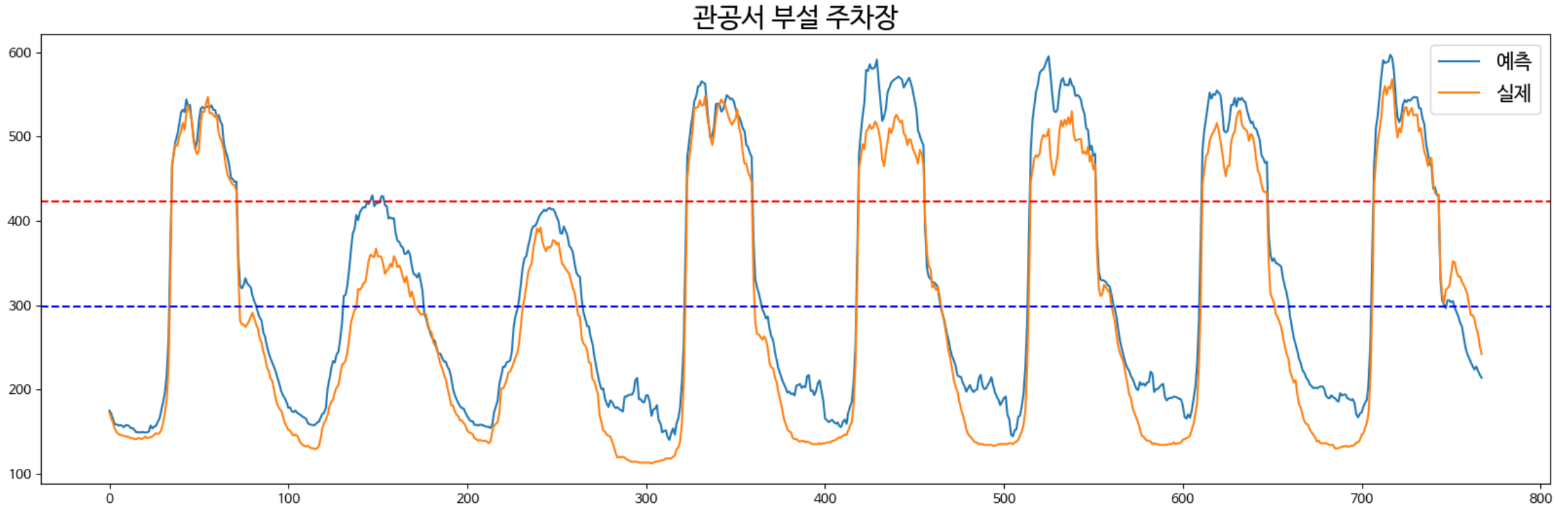
혼잡도 정의

- 혼잡도(범주형)을 타겟으로 하여 예측하는 것은 부정확하다는 판단 하,
주차대수를 수치형으로 예측 후 범주형인 혼잡도의 형태로 구간화 후 제공을 목표로 설정
- 주차장별 특징과 누적값 패턴을 반영한 혼잡도를
매우 혼잡/혼잡/보통 총 3단계로 나누었다.
- > 사용자가 알아보기 쉬울 뿐더러, 혼잡 유무가 주차장 선택에 영향을 주기 때문에 원활과 매우원활 범주는 보통에 편입시켰다.

주차장명	보통	혼잡	매우 혼잡
관공서 주차장	~ 60%	60% ~ 85%	85% ~
먹자골목 주차장	~ 60%	60% ~ 85%	85% ~
전통시장 1주차장	~ 50%	50% ~ 65%	65% ~
전통시장 2주차장	~ 50%	50% ~ 65%	65% ~



주차장별 혼잡도 기준 설정



- 데이터 EDA를 통해 주차장의 특징을 확인하여 증설 외의 다른 해답을 찾아보려 노력

서비스 제공 예시

점속 시간대를 기준으로 팝업이 출력된다
아래 예시는 12시 1분에 점속했다는 가정 하에 출력

주차장 현황

주차장1(오산시청)

현재 날짜 : 09월 18일

시간	12시 01분	12시 16분	12시 31분	12시 46분	13시 01분	14시 01분
예보						

주차장2(운암공원)

시간	12시 01분	12시 16분	12시 31분	12시 46분	13시 01분	14시 01분
예보						

주차장3(오색시장1)

시간	12시 01분	12시 16분	12시 31분	12시 46분	13시 01분	14시 01분
예보						

주차장4(오색시장2)

시간	12시 01분	12시 16분	12시 31분	12시 46분	13시 01분	14시 01분
예보						

혼잡도 예측 성능

주차장 명	구간 명	precision	recall	f1-score
관공서 주차장	매우혼잡	0.98	1.00	0.99
	weighted avg	0.97	0.97	0.97
먹자골목 주차장	매우혼잡	0.85	0.89	0.87
	weighted avg	0.97	0.97	0.97
전통시장 1주차장	매우혼잡	0.88	0.95	0.92
	weighted avg	0.92	0.92	0.92
전통시장 2주차장	매우혼잡	0.81	0.87	0.84
	weighted avg	0.91	0.91	0.91

최우선 목표인 매우 혼잡의 recall을 높이는데 주력
차 순위로 모델의 신뢰성을 해치지 않기 위해 매우혼잡의 precision도 일정수준으로 끌어올렸음
-> 그 결과, 매우 혼잡의 f1-score가 90퍼센트 전후를 기록하고 있음

기대효과

1. 오산 스마트시티

현재 오산시는 ‘함께하는 변화, 미래도시 오산’이라는 슬로건과 함께 첨단 IoT와 AI 기술을 접목한 스마트시티로의 도약을 목표로 하고 있다. 본 프로젝트는 빅데이터와 AI기술을 활용해 오산시민들의 주차수요를 파악, 예측하여 그 궤를 같이 한다. 오산시가 장기적인 목표 성장의 일부가 될 것으로 기대

2. 경제성과 교통문제 해결

대규모 공사와 예산을 대동한 물리적인 확장이 아닌 기존의 주차장의 수요를 파악하여 적은 비용으로 주차문제를 해결하고 인근 진입로의 교통체증 문제를 해결할 것으로 기대

3. 범용성

본 프로젝트와 유사하게 주변에 대규모 유인시설이 위치한 공영(사설)주차장들을 대상으로 주차면수와 수요를 예측에 적용가능

4. 인건비 감축

주차장 고도화 사업으로 시설관리공단 상황실에서 사람이 직접 CCTV를 확인해 실시간 주차장 현황을 수기로 제공하는 서비스를 개발 중

: 본 프로젝트의 모델을 적용, 서비스를 자동화한다면 연간 7500만원의 인건비를 감축할 것으로 기대

향후 연구 방향

- 1.향후 데이터의 기간이 길게 주어진다면, 빈도가 적어 주요한 요인으로 작용하지 못했던 샌드위치 데이, 월별, 공휴일, 명절 전날 등의 여부가 어떠한 영향을 끼칠 수 있을지 알 수 있을 것으로 생각된다.
- 2.주차심리 관련해서 전문가 자문을 통해 우리가 예측한 주차면수를 통해 구한 혼잡도와 입출차 혼잡도도 함께 고려한 종합적인 혼잡도를 제공할 수 있을 것이다
3. 기상 & 교통량 공공데이터 (공휴일, 주말, 연휴 대상으로 할 때 유용할 것이다)
4. 유동인구 안 쓰거를 향후연구 상관계수 (유동인구와 주차장 데이터가 일상생활 시간이라는 간섭에 너무 큰 영향을 받았다)
 - 1.상관계수 지나치게 높다
 - 2.간섭 : 전통시장과 관공서 상관계수 0.8이상 (주차장과 유동인구 사이의 관계가 아닌 주간 활동 시간대의 간섭을 크게 받는다) => (간섭 속아낼 수 없다)
5. 현재기준: 더 많은 수의 주차장
주차장 사이의 관계를 파악해 서로의 영향을 설명해보는 시도도 가능했을 듯 (실제로 오색2가 오색1을 후행하는 패턴을 보이기도 했음)

감사합니다.
