



Big Contest : 챔피언스리그

NS **SHOP+** 판매실적 예측을 통한 편성 최적화 방안 도출

팀명 : toBigCon

김미성 (팀원, pinklike23@naver.com)

김태욱 (팀장, ktu9682@gmail.com)

김태한 (팀원, kimuw42@yonsei.ac.kr)

이유진 (팀원, daebaq27@gmail.com)

목차

문제정의 &
데이터의 이해

EDA & Feature
Engineering

취급액 예측 모델

- Text CNN
- Boosting
- Hyperparameter Tuning

편성표 최적화

- 요일별 편성
- 월별 편성
- 발전 방향

● 문제정의

- EDA & Feature Engineering
- 취급액 예측 모델
- 방송 편성 최적화
- 의의 및 발전 방향

문제정의

NS Shop⁺ 판매실적 예측을 통한 편성 최적화 방안(모형) 도출

NS Shop+편성데이터(NS홈쇼핑) 를 활용하여 방송편성표에 따른 판매실적을 예측하고,
최적 수익을 고려한 요일별/ 시간대별 / 카테고리별 편성 최적화 방안(모형) 제시

- 문제정의

- EDA & Feature Engineering

- 취급액 예측 모델

- 방송 편성 최적화

- 의의 및 발전 방향

EDA & Feature Engineering



EDA

- 결측값 처리

- ① 취급액이 NaN인 경우 : 2930 / 38309 삭제
- ② 노출(분)이 NaN 인 데이터 존재 → 방송일시가 동일한 노출(분) 값으로 대체

	방송일시	노출(분)	마더코드	상품코드	상품명	상품군	판매단가	취급액
0	2019-01-01 06:00:00	20.0	100346	201072	테이트 남성 셀린리트3종	의류	39900	2099000.0
1	2019-01-01 06:00:00	NaN	100346	201079	테이트 여성 셀린리트3종	의류	39900	4371000.0
2	2019-01-01 06:20:00	20.0	100346	201072	테이트 남성 셀린리트3종	의류	39900	3262000.0

EDA

- 상품별

train set(2019-01 ~ 2019-12) 과 test set(2020-06) 에서 겹치는 비율이 적다.

→ 상품의 특성을 분류, 파악할 수 있는 feature 생성

```
print('test에서 마더코드의 unique개수', len(set(test.마더코드.unique())))
print('train에서 마더코드의 unique개수', len(set(train.마더코드.unique())))
print('train에는 있으나 test에는 없는 마더코드의 unique개수', len(set(test.마더코드.unique()) - set(train.마더코드.unique())))
print('test에는 있으나 train에는 없는 마더코드의 비율', len(set(test.마더코드.unique()) - set(train.마더코드.unique()))/len(set(test.마더코드.unique())))
```

```
test에서 마더코드의 unique개수      201
train에서 마더코드의 unique개수      687
train에는 있으나 test에는 없는 마더코드의 unique개수  116
test에는 있으나 train에는 없는 마더코드의 비율      0.5771144278606966
```

```
print('test에서 상품코드의 unique개수', len(set(test.상품코드.unique())))
print('train에서 상품코드의 unique개수', len(set(train.상품코드.unique())))
print('train에는 있으나 test에는 없는 상품코드의 unique개수', len(set(test.상품코드.unique()) - set(train.상품코드.unique())))
print('test에는 있으나 train에는 없는 상품코드의 비율', len(set(test.상품코드.unique()) - set(train.상품코드.unique()))/len(set(test.상품코드.unique())))
```

```
test에서 상품코드의 unique개수      389
train에서 상품코드의 unique개수      2031
train에는 있으나 test에는 없는 상품코드의 unique개수  365
test에는 있으나 train에는 없는 상품코드의 비율      0.9383033419023136
```


EDA

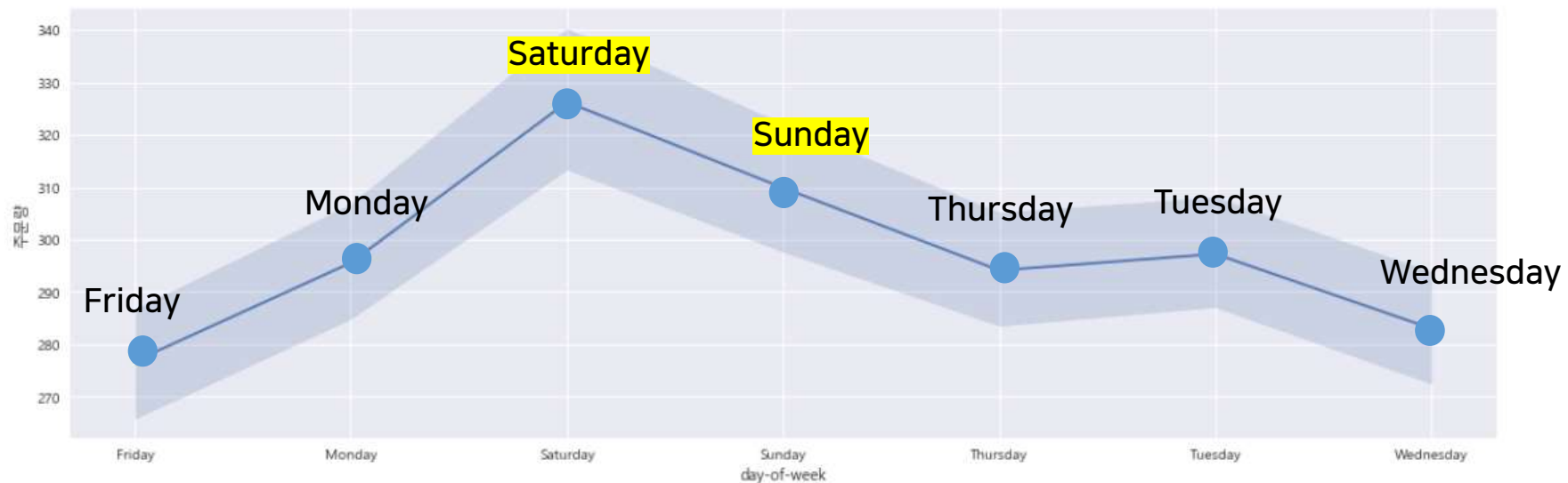
- 시간대별(요일, 월, 시간)

① 토요일과 일요일에 주문량 多

#요일별 주문량

```
sns.relplot(x="day-of-week", y="주문량", kind="line", data=df, aspect=15/5)
```

<seaborn.axisgrid.FacetGrid at 0x293a74ff0f0>



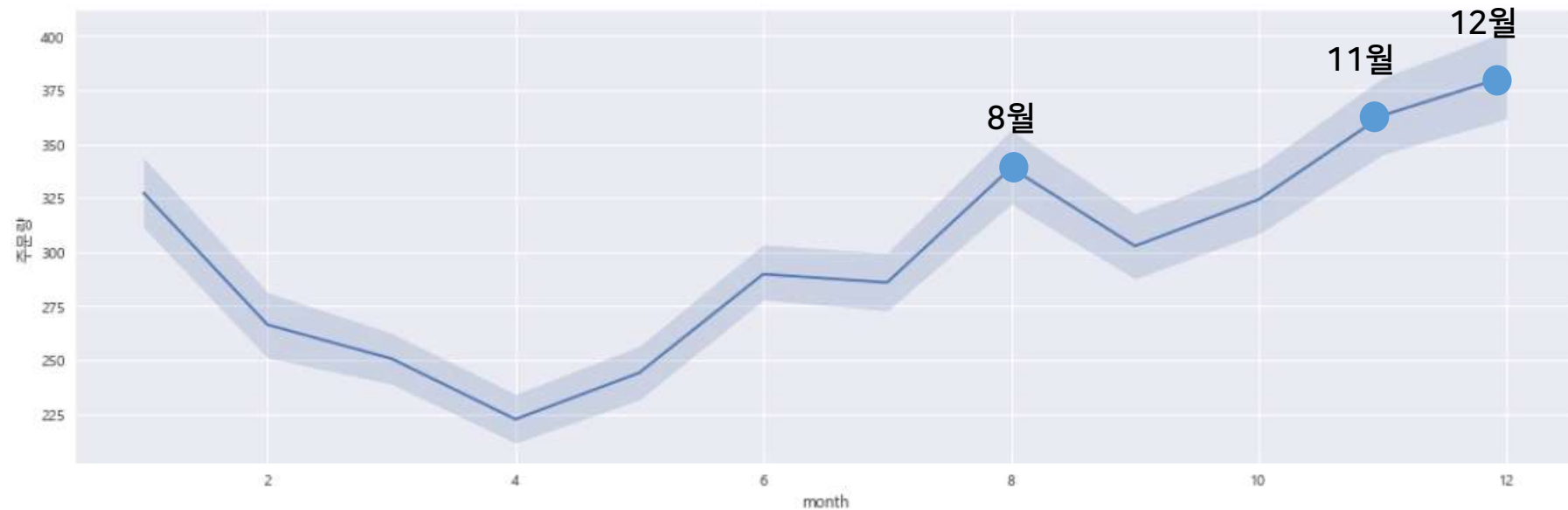
EDA

- 시간대별

② 8월, 11월, 12월에 주문량 多

```
# 월별 주문량  
sns.relplot(x="month", y="주문량", kind="line", data=df, aspect=15/5)
```

<seaborn.axisgrid.FacetGrid at 0x293a756da20>



EDA

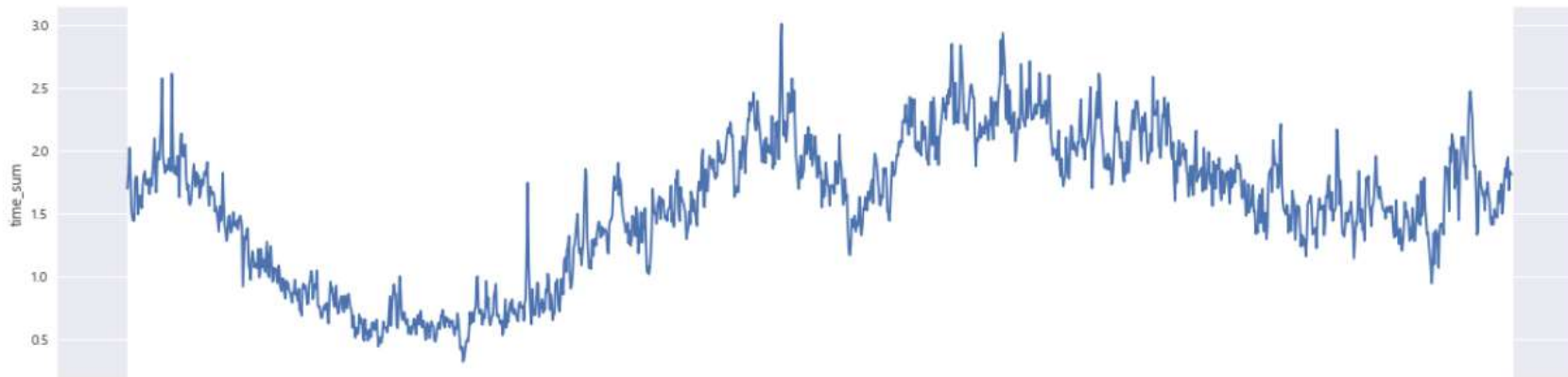
- 외부 데이터 (시청률)

: 2020년 6월은 시청률 정보가 없음.

→ 2019년 6월 요일,시간,분 별 시청률의 평균을 이용하여 2020년 6월 시청률 정보를 만들어 줌.

```
plt.figure(figsize=(20,5))  
sns.lineplot(x=temp.index, y='time_sum', data=temp)
```

<matplotlib.axes._subplots.AxesSubplot at 0x1d129e96e48>



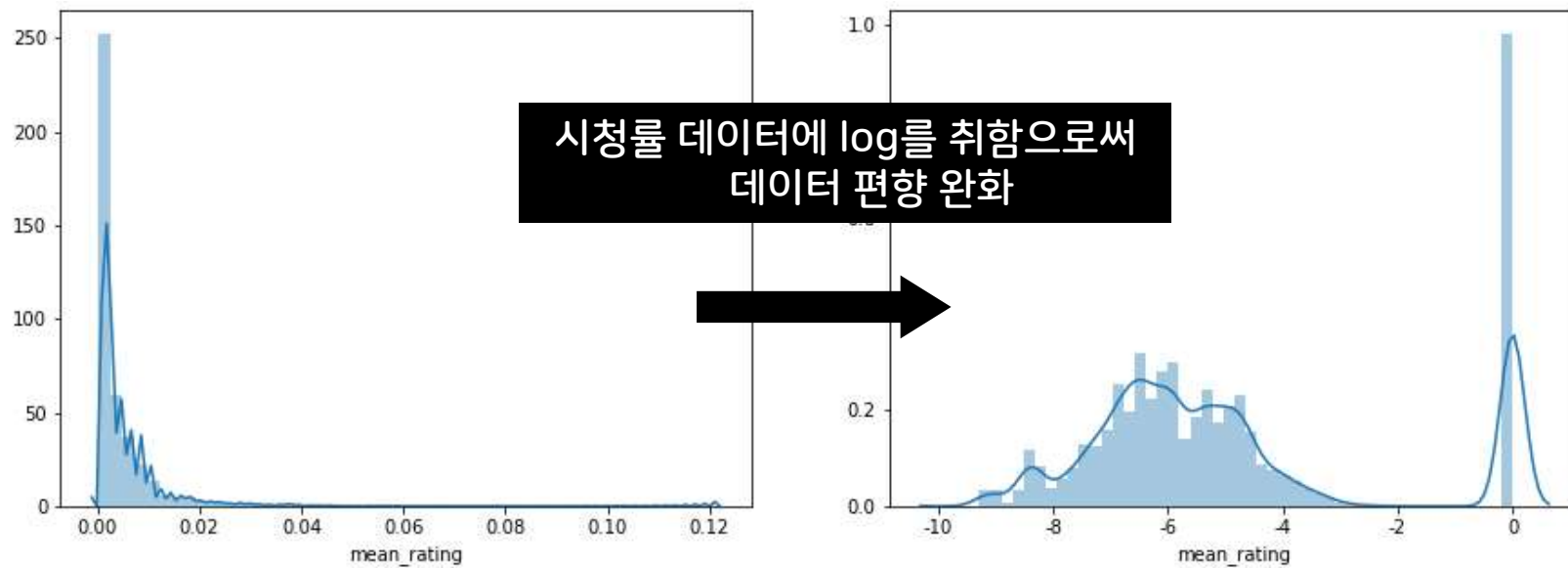
→ 2019년 시청률 분포

EDA

- 외부 데이터 (시청률)

: 2020년 6월은 시청률 정보가 없음.

→ 2019년 6월 요일,시간,분 별 시청률의 평균을 이용하여 2020년 6월 시청률 정보를 만들어 줌.



EDA

- 외부 데이터 (시청률)

: 2020년 6월은 시청률 정보가 없음.

→ 2019년 6월 요일,시간,분 별 시청률의 평균을 이용하여 2020년 6월 시청률 정보를 만들어 줌.

	방송일시	노출 (분)	마더코 드	상품코 드	상품명	상품 군	판매단 가	취급액	판매량	요일	hour	minute	date	month	season	mean_rating
20	2019-01-01 12:00:00	20.0	100088	200236	에코라믹 통주물 스톤 냄비세트	주방	60900	33823000.0	555.385878	화요일	12	0	20190101	1	겨울	0.007500
21	2019-01-01 12:20:00	20.0	100088	200236	에코라믹 통주물 스톤 냄비세트	주방	60900	59083000.0	970.164204	화요일	12	20	20190101	1	겨울	0.010000
22	2019-01-01 12:40:00	20.0	100088	200236	에코라믹 통주물 스톤 냄비세트	주방	60900	99235000.0	1629.474548	화요일	12	40	20190101	1	겨울	0.010000
23	2019-01-01 13:00:00	30.0	100305	200970	오모떼 360도 텐션업 레이스 패키지 시즌4	속옷	109000	32273000.0	296.082569	화요일	13	0	20190101	1	겨울	0.010000
24	2019-01-01 13:30:00	30.0	100305	200970	오모떼 360도 텐션업 레이스 패키지 시즌4	속옷	109000	79273000.0	727.275229	화요일	13	30	20190101	1	겨울	0.010000
25	2019-01-01 14:00:00	30.0	100377	201220	그렉노먼 남성 구스다운 통 벤치코트	의류	119000	11269000.0	94.697479	화요일	14	0	20190101	1	겨울	0.005900
26	2019-01-01 14:00:00	30.0	100377	201226	그렉노먼 여성 구스다운 통 벤치코트	의류	119000	20841000.0	175.134454	화요일	14	0	20190101	1	겨울	0.005900

EDA & Feature Engineering

상품명에서 추출

시간 정보 추출

상품군에서 추출

취급액에서 추출

외부 데이터

상품명에서 추출

- ① 중분류 / 소분류
- ② Muil : 무이자, 일시불 여부
- ③ Brand : 브랜드 별로 구분
- ④ Season_prod : 계절성 상품 분류
- ⑤ 남/여 : 남성용품, 여성용품 구분
- ⑥ Prod_size : 상품의 크기

ex) 소들넉 소 갈비탕 → 국

ex) 무이자 LG 통돌이 세탁기 → 무

ex) 오모떼 360도 텐션업 레이스 → 오모떼

ex) 에어컨 → 여름가전

ex) 원피스 → 여, 면도기 → 남

ex) 쿠푸밥솥 6인 → 6

```
print('test에서 중분류의 unique개수', len(set(test.중분류.unique())))
print('train에서 중분류의 unique개수', len(set(train.중분류.unique())))
print('train에는 있으나 test에는 없는 중분류의 unique개수', len(set(test.중분류.unique()) - set(train.중분류.unique())))
print('test에는 있으나 train에는 없는 중분류의 비율', len(set(test.중분류.unique()) - set(train.중분류.unique()))/len(set(test.중분류.unique())))
```

test에서 중분류의 unique개수	57
train에서 중분류의 unique개수	89
train에는 있으나 test에는 없는 중분류의 unique개수	0
test에는 있으나 train에는 없는 중분류의 비율	0.0

→ 상품명에서 최대한 많은 정보를 추출하여 train set 에 없는 상품을 test 에서 예측하기 쉽게 함.

EDA & Feature Engineering

상품명에서 추출

시간 정보 추출

상품군에서 추출

취급액에서 추출

외부 데이터

['스웨터' '브라' '팬츠' '외투' '넴비' '런닝' '고기' '수산물' '농산물' '화장품' '세탁기' '소도구' '비타민'
'인덕션' '호빵' '오븐' '팬티' '냉장고' 'TV' '레깅스' '신발' '백' '블렌더' '가전' '쥬스' '중도구' '란주'
'쥬얼리' '삼푸' '티셔츠' '견인기기' '노트북' '과일' '건조대' '백팩' '도마' '티포트' '언더셔츠' '공기청정기'
'건조기' '모자' '밍크' '기타_농수축' '침대' '싱크선반' '뿌리뽑' '가구' '거실장' '트레이닝' '소파' '선글라스'
'청소기' '에어컨' '유산균' '하이바스' '베이지' '매트' '스타일러' '패키지' '악어' '썸머' '침구' '고대기'
'김치' '여행가방' '커튼' '카페트' '소가죽' '주니어' '환' '블라우스' '시공패키지' '사철' '다이어트' '원피스'
'대자리' '양산/우산' '선풍기' '김치냉장고' '도어락' '협탁' '홈웨어' '지갑' '가마솥' '세트' '리프트' '브러쉬'
'벽지' '기모']

▶ 중분류

돌다아님 43171
일시불 7797 ▶ muli
무이자 7751
Name: muli, dtype: int64

0 28253
2 3850
4 2945
1 1504
3 1260
5 271 ▶ prod_size
6 12
Name: prod_size, dtype: int64

all 45620
여름의류 4385
여름가전 4027
겨울의류 2461
여름잡화 1023
겨울가전 674
여름식품 352
겨울식품 132 ▶ season_prod
겨울잡화 45
Name: season_prod, dtype: int64

all 44976
여성 7978 ▶ 남/여
남성 5765
Name: 남여, dtype: int64

['니트' '브라' '팬츠' '무스탕' '넴비' '레이스' '코트' '고기' '오징어' '물' '피부세드' '모아' '세탁기'
'티플렉' '옷걸이' '루테인' '채집' '법술' '가스레인지' '황태' '고등어' '간식' '꽃게' '에어프라이어' '대구'
'드로즈' '런닝' '냉장고' 'TV_55' 'TV_65' 'TV_70' '레깅스' '부츠' '다시락' '크림' '백' '익서'
'침방' '종이' '트렁크' '배대' '알배추전역' '리빙박스' '차비' '사계' '트러드먼트' '핀터' '베스트' '티셔츠'
'공창' '비타민' '립스틱' '공구' '스웨터' '팬티' '후라이팬' '꽃' '견인기기' '에어쿠일' '패드' '해수' '노트북'
'다이아몬드' '커피머신' '오븐' '과일' '두유' '진공포장기' '통배프' '백팩' '멀티과' '참소쿠리' '밴드백' '용기'
'그릴' 'TV_75' '종삼' '마스크' '옥수수' '갈치' '래민지' '언더셔츠' '세이버' '공기청정기' '건조기' '문어'
'원포' '미니건조기' '피에이케' '모자' '밍크' '선식' '도마' '매생이' '도트백' '행우' '침대_상글' '침대_흰'
'침대_립' '란주' '오메가3' '카타' '간강기능' '싱크선반' '욕류' '염색' '워킹화' '세탁' '시래기' '행거' '외자'
'행소기' '자세밴드' '오렌지' '슬립온' '이불보' '블랙아일랜드' '가디건' '국' '레드벤' '드레킹화' '드레킹'
'리블라이너' '슈즈' '심유유연제' '클러너' '선글라스' '크로스' '스쿼트머신' '스니커즈' '에어컨_홈터'
'에어컨_스탠드' '벤치' '바이오' '수트' '하이바스' '입블러' '선반' '남자' '베이스레이어' '코일카세트' '언더팔'
'왕조함' '원드' '침화' '세트' '스타일러' '매시지' '쿠션' '패키지' '악어' '물' '기타_농수축' '침구_슈퍼싱글'
'침구_흰' '침구_립' '모시백' '고대기' '김치' '통침대_상글' '통침대_흰' '여행가방' '커튼_중형' '커튼_슈퍼특대형'
'커튼_대형' '커튼_특대형' '카세트_중형' '카세트_대형' '카세트_특대형' '맛들' '레모나' '세안' '가자미' '알상불'
'소가죽' '소파_4인용' '소파_6인용' '소파_3인용' '버킷백' '백걸이_B6' '레닝' '주니어' '글라스'
'통함' '폴렌다' '블라우스' '서클레이터' '원피스' '시공패키지' '행마' '아미스' '장어' '선물' '직접기'
'백걸이_07' '빨래' '커튼' '리스틀릿' '사철' '다이어트' '산통기' '러닝화' '셀레일' '배관청소' '수' '사워기'
'매일' '슈퍼링' '상계함' '대트릭스' '백스쿠터' '건조대' '방어' '청감상어' '에센스' '대자리'
'에어컨_홈터_클래식' '에어컨_스탠드_클래식' '양산' '수세미' '목걸이' '팔찌' '위터' '민자' '감말행어' '에초기'
'이비스커스' '지퍼백' '도카니탈' '홈터' '귀걸이' '유산균' '패드' '기타_잡화' '브로우' '연도기' '연도날'
'달팔' '상' '스카프' '고춧가루' '도마_흰' '민마' '자마살' '강정배' '블루투스' '차막' '모나코샤늘' '패죽순'
'낙지' '블랙박스' '백도' '오일' '식기세척기' '채지' '벤신족' '아이라이너' '김치냉장고' '아이스타일러' '침프스'
'도어락' '다리미' '세탁장' '배리칼' '비드줄' '랜투엔' '다지기' '립탁' '세드' '거실장' '홈웨어' '지갑'
'머플러' '가마솥' '젓갈' '구스다운' '분통' '문수레트' 'LED' '축재조기' '카우치소파' '리프트' '엔진스틱'
'조기' '오디오' '골드바' '브러쉬' '이치웨세드' '기모' '다퓨저' '어너웨이' '도마_슈퍼싱글' '도마_흰' '히터'
'카세트_정복대형' '단열벽지' '마스크팩' '세우' '속눈썹' '꽃게' '올라겐' '피온대이선' '충전기' '팔팔게' '세게'
'테라피' '워킹머신' '배준' '티포트' '굴비' '벽지' '파장대' '방한화' '카우치침대' '장금' '사워기'
'과해기' '치킨' '레이블' '합탈' '플래티늄' '굴' '캐리올' '술다' '업스티']

▶ 소분류

EDA & Feature Engineering

상품명에서 추출

시간 정보 추출

상품군에서 추출

취급액에서 추출

외부 데이터

시간 정보 추출

- ① Weekend : 주말 여부
- ② Group : 같은 상품의 노출(분) 순서
- ③ Prime time : 점심, 오후, 새벽, 저녁, no_prime 으로 특정시간 구분
- ④ Holiday_yn : 공휴일 유무
- ⑤ Holiday : 공휴일 종류

	방송일시	노출(분)	마더코드	상품코드	상품명	상품군	판매단가	취급액	group
0	2019-08-26 00:20:00	20.0	100000	200000	엘로엘 아쿠아클린 마스크	이미용	79900	12078000.0	1
1	2019-08-26 00:40:00	20.0	100000	200000	엘로엘 아쿠아클린 마스크	이미용	79900	17338000.0	2
2	2019-08-26 01:00:00	20.0	100000	200000	엘로엘 아쿠아클린 마스크	이미용	79900	18195000.0	3

EDA & Feature Engineering

상품명에서 추출

시간 정보 추출

상품군에서 추출

취급액에서 추출

외부 데이터

상품군에서 추출

- ① Expensive : 상품군 중 취급액이 높은 상품군 분류 (상위 80%)
- ② Mid : 상품군 중 취급액이 중간인 상품군 분류
- ③ Cheap : 상품군 중 취급액이 낮은 상품군 분류 (하위 20%)

[각 상품군 전체 개수 중 상위 80프로안에 들어가는 개수의 비율]

의류 0.24295832647010376
속옷 0.2636672325976231
주방 0.1999757605138771
농수축 0.3861125062782521
이미용 0.3423312883435583
가전 0.13260298984366084
생활용품 0.1099375900144023
건강기능 0.32401656314699795
잡화 0.10799504183452123
가구 0.06704601682335477
침구 0.014778325123152709

→ 농수축은 30% 가 취급액이 높은 상품에 속한다.

```
def expensive(x):  
    if x in ['농수축', '이미용']:  
        return 1  
    else:  
        return 0  
performance_data_2['expensive'] = performance_data_2['상품군'].apply(lambda x: expensive(x))  
performance_data['expensive'] = performance_data['중분류'].apply(lambda x: expensive(x))
```

```
def mid(x):  
    if x in ['의류', '건강기능', '속옷']:  
        return 1  
    else:  
        return 0  
performance_data_2['mid'] = performance_data_2['상품군'].apply(lambda x: mid(x))  
performance_data['mid'] = performance_data['상품군'].apply(lambda x: mid(x))
```

```
def cheap0(x):  
    if x in ['가구', '침구']:  
        return 1  
    else:  
        return 0  
performance_data_2['cheap0'] = performance_data_2['상품군'].apply(lambda x: cheap0(x))  
performance_data['cheap0'] = performance_data['상품군'].apply(lambda x: cheap0(x))
```

EDA & Feature Engineering

상품명에서 추출

시간 정보 추출

상품군에서 추출

취급액에서 추출

외부 데이터

취급액에서 추출

- ① 취급액_mean_quantile_10 : 중분류 및 소분류 기준으로 취급액 평균 10분위 구분
- ② 취급액_mean_log_quantile_10 : 중분류 및 소분류 기준으로 취급액 log 평균 10분위 구분

	취급액_mean_log_quantile_10	취급액_mean_quantile_10
20	3	3
21	3	3
22	3	3
23	5	6
24	5	6
25	7	7
26	7	7
27	7	7
28	7	7
29	9	9
30	9	9

→ log 를 씌운 분포와 서로 차이를 보임.

EDA & Feature Engineering

상품명에서 추출

시간 정보 추출

상품군에서 추출

취급액에서 추출

외부 데이터

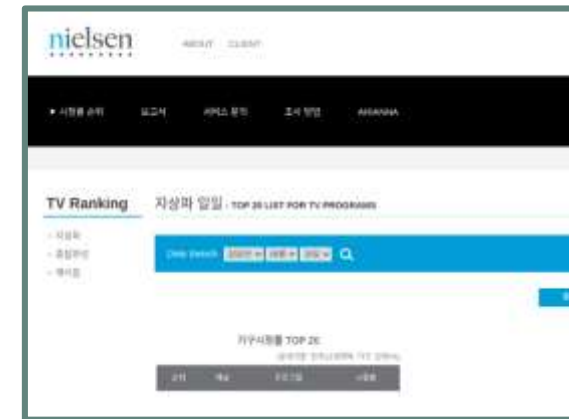
외부 데이터

① 주가 데이터 : 주가 데이터를 통한 시장 정보 반영

② 날씨 데이터 : 기온, 습도, 풍속, 강수량 정보 반영

③ 뉴스 시청률 : KBS 9시 뉴스 시청률 정보 반영

→ 외부 데이터를 활용하여 코로나 등 소비에 영향을 미치는 외부요인을 반영하려 노력.



market : <http://www.krx.co.kr/main/main.jsp>

weather : https://data.kma.go.kr/cmmn/main.do;jsessionid=0fGAqAErnuTWTCjBKbCuIrrMHudZZO1LC7F2Kp6BL8Xtsnm5V9UTLzFLm9JedHuI.was02_servlet_engine5

news rating : https://www.nielsenkorea.co.kr/tv_terrestrial_day.asp?menu=Tit_1&sub_menu=1_1&area=00&begin_date=

EDA & Feature Engineering

[최종 Feature]

#	Column	Non-Null Count	Dtype
0	방송일시	58719 non-null	datetime64[ns]
1	노출(분)	58719 non-null	float64
2	마더코드	58719 non-null	object
3	상품코드	58719 non-null	object
4	상품명	58719 non-null	object
5	상품군	58719 non-null	object
6	판매단가	58719 non-null	float64
7	취급액	58719 non-null	float64
8	판매량	58719 non-null	float64
9	요일	58719 non-null	object
10	hour	58719 non-null	int64
11	minute	58719 non-null	int64
12	date	58719 non-null	object
13	month	58719 non-null	int64
14	season	58719 non-null	object
15	mean_rating	58719 non-null	float64
16	holiday	58719 non-null	object
17	holiday_yn	58719 non-null	float64
18	기온	58719 non-null	float64
19	누적강수량	58719 non-null	float64
20	풍속	58719 non-null	float64
21	습도	58719 non-null	float64
22	현재지수	58719 non-null	float64
23	대비	58719 non-null	float64
24	등락률(%)	58719 non-null	float64
25	배당수익률(%)	58719 non-null	float64
26	주가이익비율	58719 non-null	float64
27	주가자산비율	58719 non-null	float64
28	시가지수	58719 non-null	float64
29	고가지수	58719 non-null	float64
30	저가지수	58719 non-null	float64
31	거래량(천주)	58719 non-null	float64
32	거래대금(백만원)	58719 non-null	float64
33	상장시가총액(백만원)	58719 non-null	float64
34	뉴스시청률	58719 non-null	float64
35	prime_time	58719 non-null	object
36	소분류	58719 non-null	object
37	중분류	58719 non-null	object
38	브랜드	58719 non-null	object
39	group	58719 non-null	int64
40	남여	58719 non-null	object
41	mul	58719 non-null	object
42	season_prod	58719 non-null	object
43	취급액_mean_quantile_10	58719 non-null	int64
44	취급액_mean_log_quantile_10	58719 non-null	int64
45	취급액_mean_quantile_10_소분류	58719 non-null	int64
46	취급액_mean_log_quantile_10_소분류	58719 non-null	int64
47	expensive	58719 non-null	int64
48	mid	58719 non-null	int64
49	cheap0	58719 non-null	int64
50	weekend	58719 non-null	int64
51	middle_class	58719 non-null	int64
52	prod_size	58719 non-null	int64

- 문제정의
- EDA & Feature Engineering
- **취급액 예측 모델**
- 방송 편성 최적화
- 의의 및 발전 방향

취급액 예측 모델

	Model1	Model2	Model3
Dataset	Dataset selection <u>상품군 별 dataset 분리</u>	Dataset selection <u>상품군 별 dataset 분리 X</u>	시간 관련 변수, 상품명
방식	LGBM, XGB 앙상블	LGBM, XGB, Catboost 앙상블	TextCNN

Weighted sum

Weighted sum

Weighted sum

취급액 예측 모델

TextCNN

취급액 예측 모델

TextCNN

(도넛)일시불 쿠키전기밥솥 10인용

- 무이자/일시불, 상품특성, 브랜드, 상품의 용량, 크기, 용도 등 상품명에 담긴 정보 多
- 기존의Tree based 모델(LGBM, XGBoost 등)은 이런 정보를 완벽히 활용하기 어려움.

∴ 상품명이 가지고 있는 text 정보를 취급액 예측에 바로 적용할 수 있는 모델 필요

취급액 예측 모델

TextCNN: 전처리 및 토큰나이징

상품명 예시

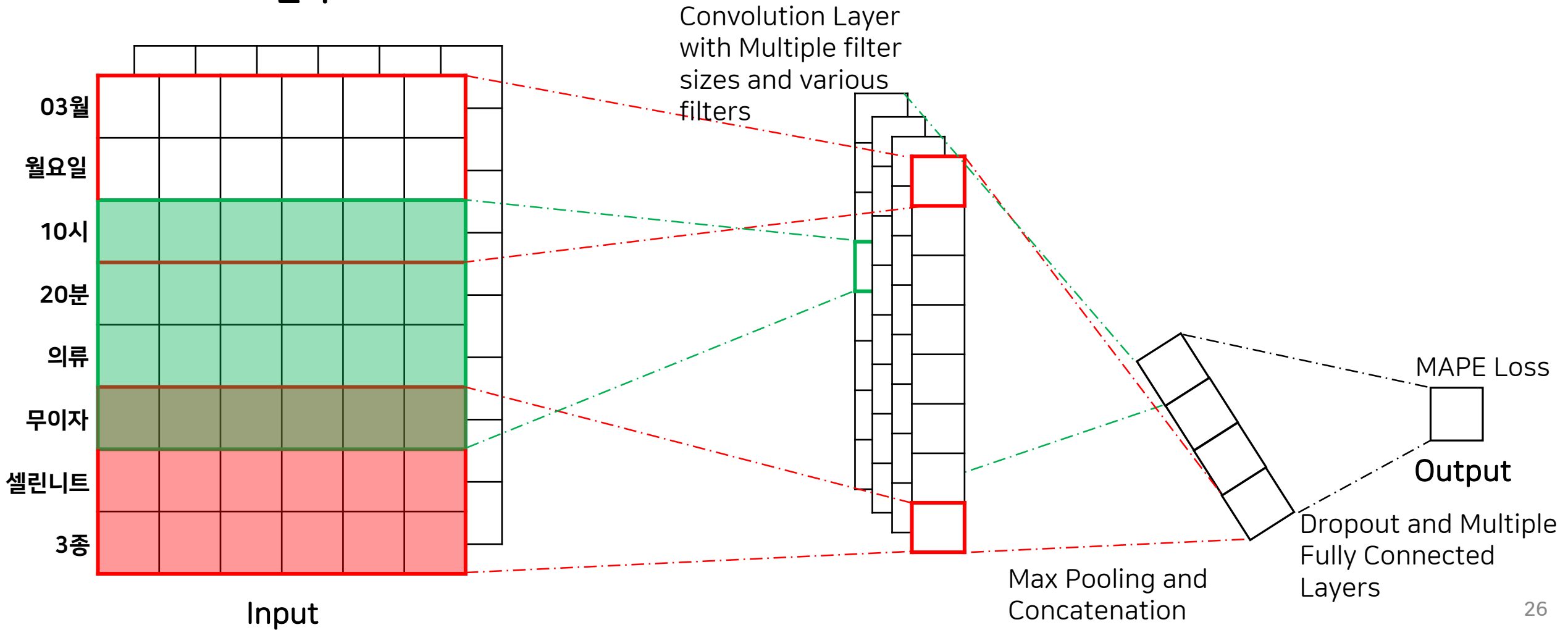
- 프라다 레터링 카메라백
- 레이프릴 데일리 코튼 보정팬티 12종
- 휴렉 음식물처리기
- 생줄랑 18k 코코 천연 다이아 팔찌
- 피시원 국내산 손질 갑오징어 9팩
- 무이자 린나이 간편쿡 자동불꽃조절 가스레인지 2구

- 기존 텍스트 시퀀스와 달리 조사나 연결 어미가 적고 **term 단위**로 의미가 chunking 되어있음
- 브랜드 등의 **고유명사**가 빈번히 등장

→ 띄어쓰기, 오타 교정 위주의 전처리 및 term 단위의 의미를 해치지 않도록 space를 기준으로 split

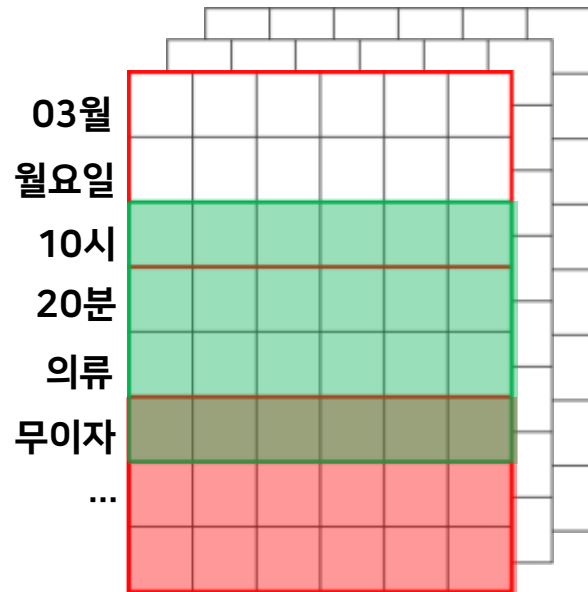
취급액 예측 모델

TextCNN: 모델 구조



취급액 예측 모델

TextCNN: input



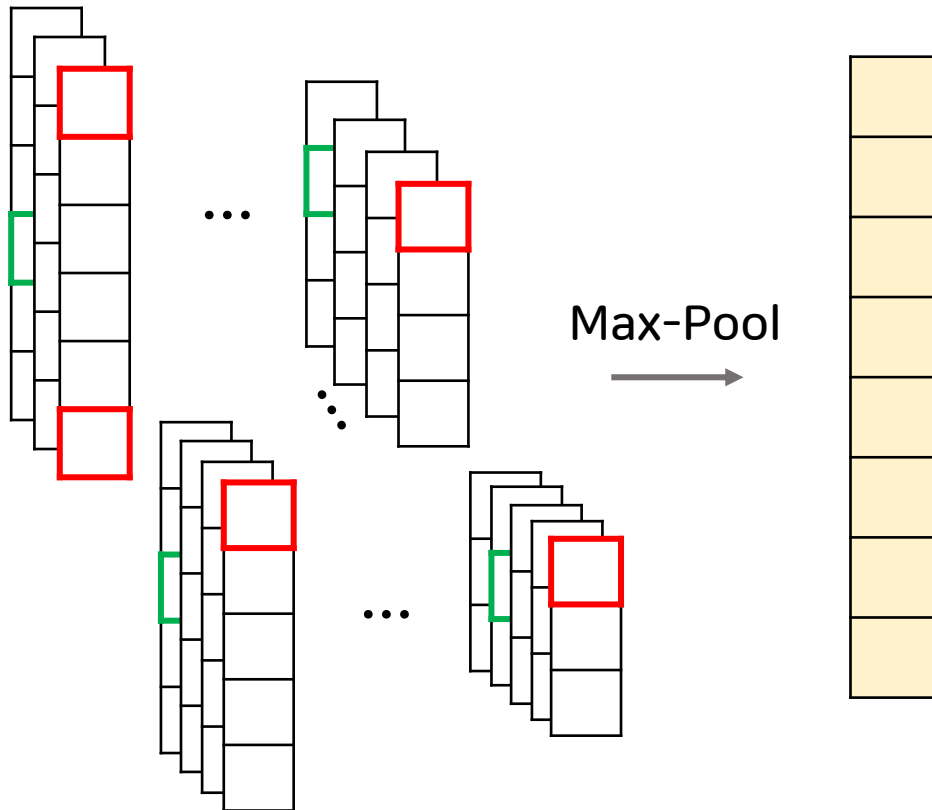
2019-01-03 15:00	20	100837	202455	(도넬)무이자 쿠키전기밥솥 10인용	주방	218,000	13,222,000
2019-01-03 15:00		100837	202458	(도넬)일시불 쿠키전기밥솥 10인용	주방	208,000	9,656,000
2019-01-03 15:00		100837	202456	(도넬)무이자 쿠키전기밥솥 6인용	주방	208,000	1,530,000
2019-01-03 15:00		100837	202457	(도넬)일시불 쿠키전기밥솥 6인용	주방	198,000	4,272,000

2019-01-08 18:00	20	100448	202098	일시불 쿠키 풀스텐 압력밥솥 10인용 (A1)	주방	158,000	13,339,000
2019-01-08 18:00		100448	202093	무이자 쿠키 풀스텐 압력밥솥 10인용(A1)	주방	168,000	4,763,000
2019-01-08 18:00		100448	202100	일시불 쿠키 풀스텐 압력밥솥 6인용(A1)	주방	148,000	3,669,000
2019-01-08 18:00		100448	202095	무이자 쿠키 풀스텐 압력밥솥 6인용(A1)	주방	158,000	3,368,000

1. 같은 상품인 경우에도 시간대에 따라 매출이 달라짐
2. 새로운 데이터가 들어올 때 OOV(Out of Vocabulary) 문제 발생 가능
→ 시간, 상품군, 상품 소분류 등 상품명에 추가하여 모델 input 생성

취급액 예측 모델

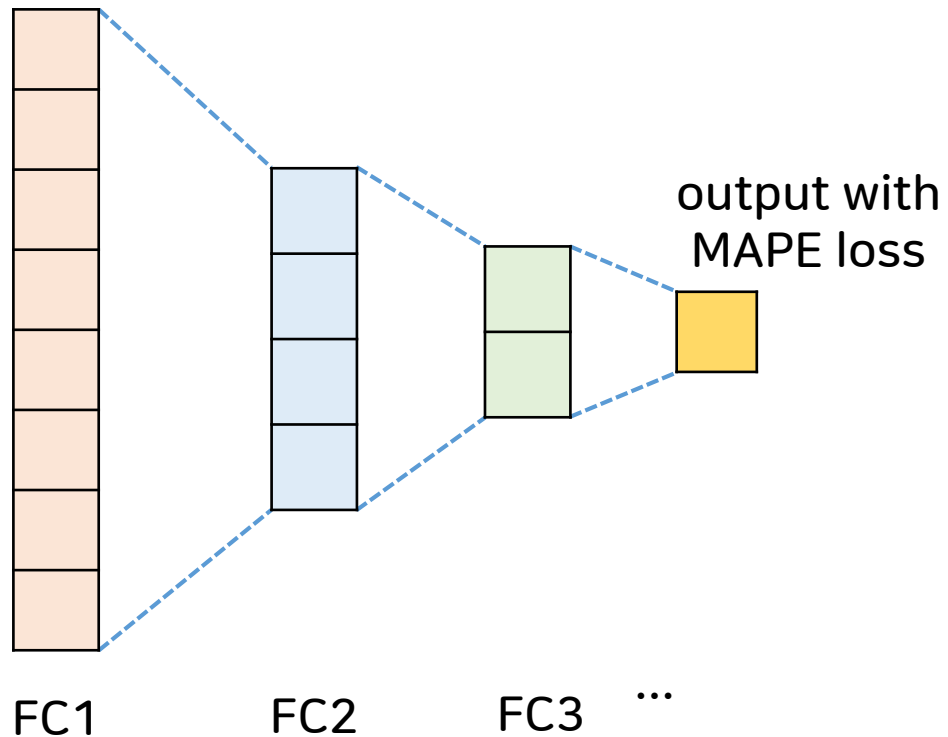
TextCNN: Convolution and MaxPooling



1. 하나의 conv layer에서 9개의 독립적인 convolution (window size가 context를 파악 → 다양한 window size 이용)
2. window size마다 다른 필터 개수: 고려할 문맥 범위에 가중치 부여 효과
(1부터 10까지 다양한 window size, 15 ~ 35까지 다양한 필터 개수 사용)
3. max pooling layer: 상품명에서 가장 두드러지는 특징을 추출

취급액 예측 모델

TextCNN: Multiple Fully Connected Layers



1. 여러 개의 linear layer를 통과
→ network가 표현 가능한 분포의 다양화
→ 취급액에 대한 예측력을 높이고자 함.
2. MAPE Loss
→ 편향된 취급액의 분포를 상대적으로
원활히 학습하기 위해 MAPE를 사용

취급액 예측 모델

Tree Model

- LGBM, XGB, CAT

취급액 예측 모델

- LGBM (Light Gradient Boosting Model)

- 회귀 트리모델
- 트리의 균형을 맞추지 않고, leaf 노드를 지속적으로 분할하면서 진행.
- 가볍고 속도가 빠름.



- XGBoost (eXtreme Gradient Boosting)

- 회귀 트리모델
- 분류와 회귀 모델에서 뛰어난 성능 발휘

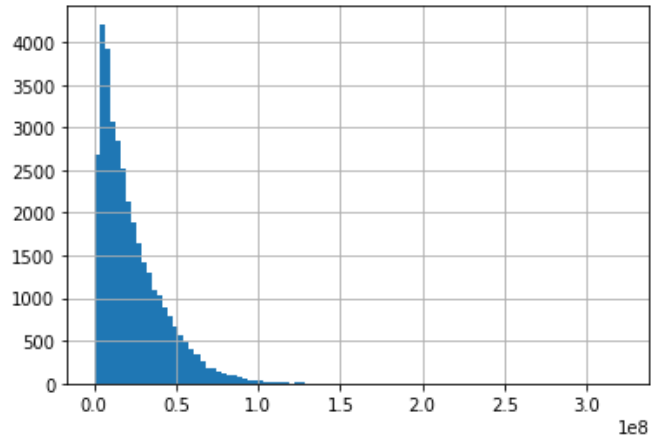
- CatBoost (Categorical Boosting)

- 상대적으로 범주형(categorical) 변수를 처리하는 데 유용함
- 기존 범주형 변수를 수치형 변수로 변환하는 방법의 overfitting의 문제 보완

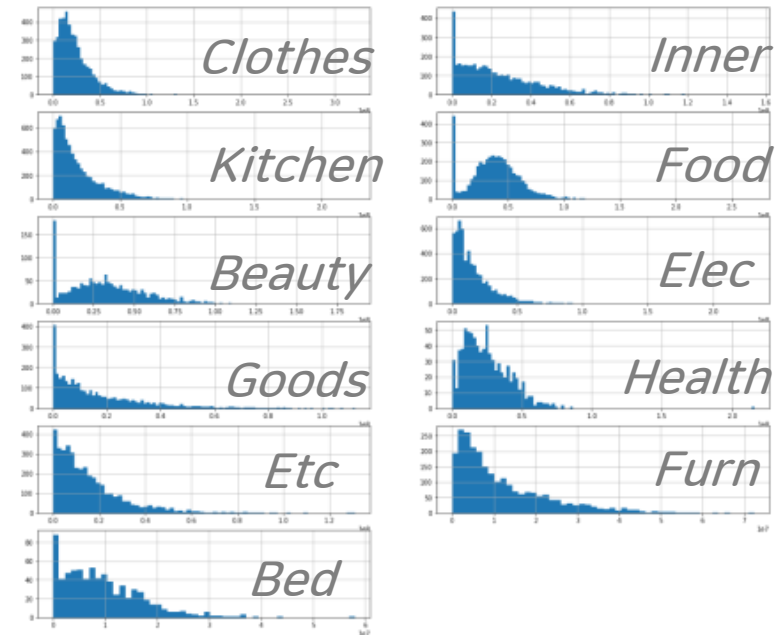


취급액 예측 모델

상품군별 모델 학습



[전체 취급액 분포]

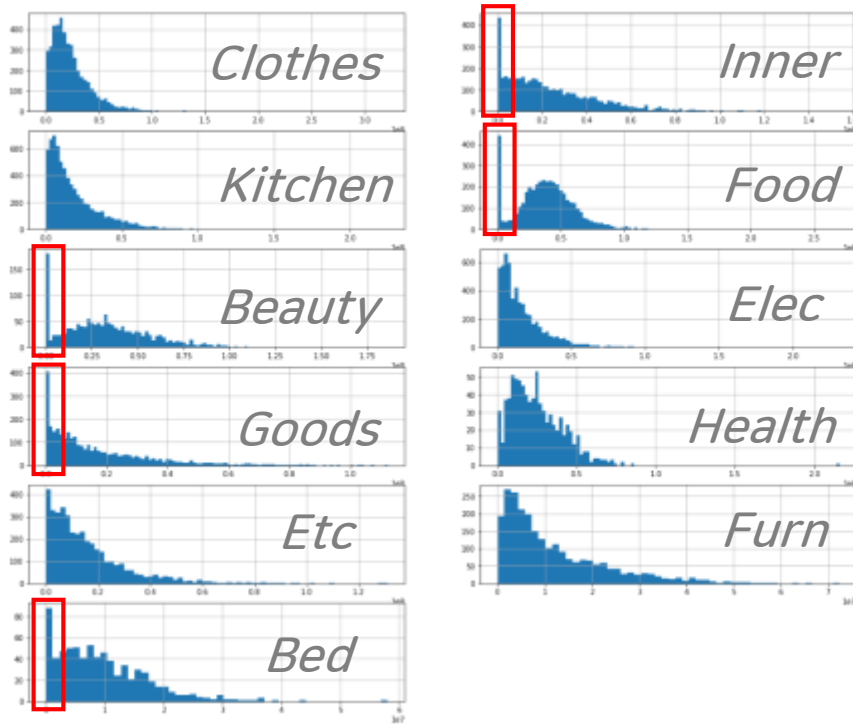


[상품군별 취급액 분포]

전체 취급액 분포와 상품군별 취급액 분포가 전반적으로 상이함 → **상품군별 모델**의 필요성

취급액 예측 모델

상품군별 모델 학습



[상품군별 취급액 분포]

감마분포와 유사한 형태

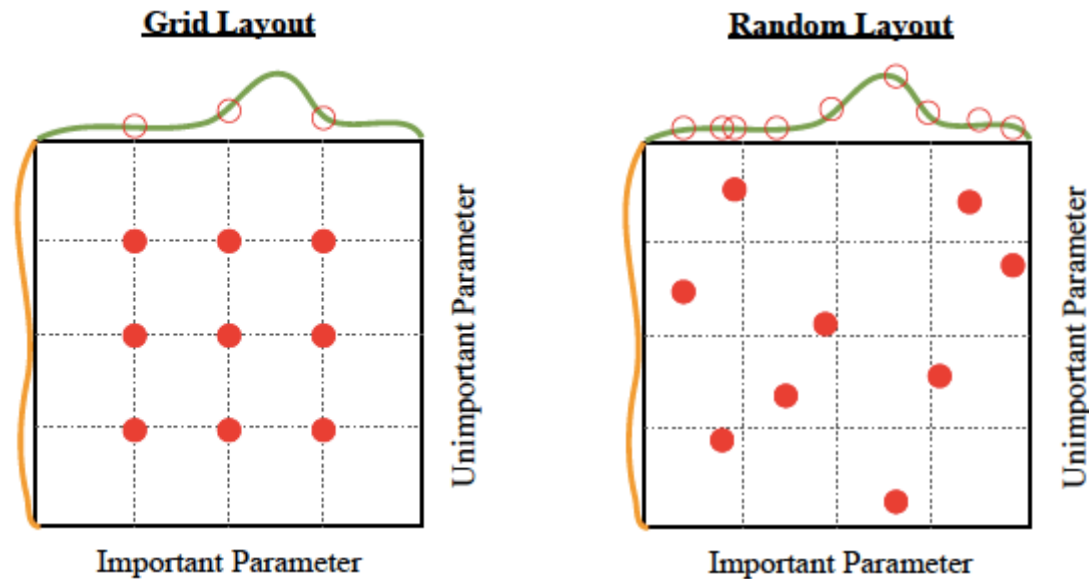
몇몇 상품군(속옷, 농수축, 이미용, 잡화, 침구 등)에 대해서는 0에 가까운 값으로 뭉쳐 있는 경향 0

∴ Regression 트리모델의 목적함수를

'tweedie' 등으로 조정함으로써 취급액의 분포를 반영

취급액 예측 모델

Hyperparameter Tuning



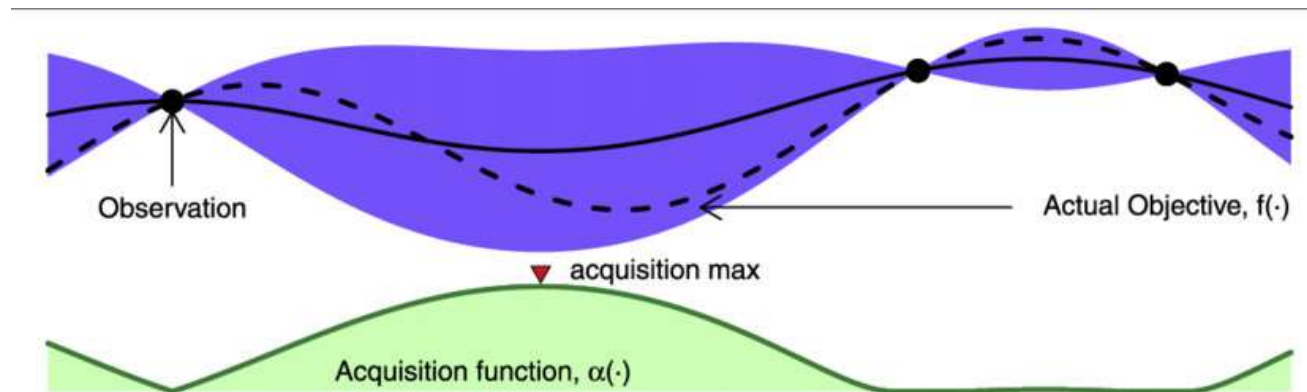
기존의 방법론은 이전 step에서의 정보를
기억하지 못하고 처음부터 다시 search

→ 비효율적

취급액 예측 모델

Hyperparameter Tuning: Bayesian Optimization

- Hyperparameter 선택하는 과정을 gaussian process를 이용하여 mean, variance 2개의 density function으로 근사하는 최적화 방법
- **이전 파라미터의 정보**가 variance의 kernel function 계산 과정에서 **반영**되어 남아있음



취급액 예측 모델

Hyperparameter Tuning: 예시

```
# all
bo = BayesianOptimization(lgb_eval, pbounds = params, random_state=0)
bo.maximize(n_iter = 100, init_points = 20, acq = 'ei', xi = 0.01, random_state = 0)
```

```
LGBMRegressor(**make_param(bo))
```

```
LGBMRegressor(learning_rate=0.2, max_depth=20, min_child_samples=35,
               min_data_in_leaf=333, min_split_gain=0.001, n_estimators=906,
               num_leaves=150, objective='tweedie', sub_sample=0.7)
```

iter	target	learn...	max_depth	min_ch...	min_ch...	min_da...	min_sp...	n_esti...	num_le...	sub_sa...
1	-44.3	0.1323	16.3	60.28	54.49	423.7	0.06494	1.156e+0	914.5	0.9891
2	-45.28	0.1075	17.29	52.89	56.8	925.6	0.008033	630.7	32.46	0.9498
3	-44.93	0.1667	18.31	97.86	79.92	461.5	0.07827	677.4	659.6	0.743
4	-44.32	0.1917	13.78	41.47	26.46	774.2	0.04616	1.353e+0	31.02	0.8853
5	-42.31	0.1418	15.02	94.37	68.18	359.5	0.04427	1.546e+0	72.95	0.9
6	-44.4	0.1506	9.735	12.89	31.54	363.7	0.05745	1.158e+0	1.012e+0	0.7306
7	-43.01	0.08133	9.097	65.31	25.33	466.3	0.0252	738.5	123.7	0.8969
8	-44.36	0.07073	9.556	36.87	82.1	97.1	0.08396	644.1	1e+03	0.8406
9	-37.8	0.1965	14.86	73.93	3.92	282.8	0.0129	944.2	132.2	0.7954
10	-44.23	0.1121	7.834	69.25	56.66	265.4	0.0528	640.9	594.9	0.9788
11	-39.59	0.09779	15.68	13.18	71.63	289.4	0.01914	1.38e+03	32.35	0.9487
12	-45.76	0.0507	15.81	27.0	73.52	962.2	0.02563	1.364e+0	611.1	0.8717
13	-41.7	0.08346	19.39	44.71	84.64	699.5	0.03045	1.721e+0	413.3	0.9643
14	-46.84	0.1372	18.46	69.25	72.53	501.3	0.09565	1.466e+0	440.9	0.8819
15	-47.54	0.05288	10.92	66.02	29.01	618.0	0.04345	703.2	313.9	0.871
16	-46.56	0.1386	14.47	65.32	65.21	431.4	0.08976	1.051e+0	453.1	0.9676
17	-48.1	0.1709	16.15	10.02	91.95	714.2	0.09989	724.2	890.5	0.7487
18	-44.39	0.1423	8.61	84.8	80.73	569.1	0.04131	603.8	717.8	0.8361
19	-39.8	0.1583	18.26	97.55	85.58	11.71	0.03664	1.595e+0	185.7	0.8563
20	-41.84	0.05815	9.6	1.852	79.37	223.9	0.03519	1.892e+0	724.9	0.7096
21	-36.61	0.1945	14.93	62.33	9.066	225.4	0.008308	1.017e+0	51.42	0.8094
22	-35.35	0.2	17.99	72.43	0.001	92.16	0.001	1.001e+0	57.62	0.7
23	-39.52	0.05	20.0	0.0	0.001	0.0	0.001	1.093e+0	17.98	1.0
24	-35.46	0.2	17.16	100.0	0.001	103.7	0.001	913.7	65.69	0.7
25	-35.78	0.2	20.0	100.0	0.001	104.7	0.001	993.8	117.4	0.7
26	-42.21	0.1738	7.154	94.26	82.18	119.3	0.04653	972.7	57.88	0.9302
27	-35.77	0.2	20.0	28.16	0.001	121.9	0.001	951.8	87.54	0.7
28	-36.24	0.2	20.0	80.26	0.001	27.96	0.001	931.2	82.16	0.7
29	-39.69	0.1862	11.96	87.91	2.012	218.7	0.02417	1.164e+0	134.5	0.9551

취급액 예측 모델

Feature selection

- Permutation importance
: 해당 변수에 noise를 주고(shuffling)
모델 학습 시, 모델 성능이 올라가면
변수 중요도가 ↓

중요한 feature

성능이 올라간 feature (안 좋은 feature)

영향을 미치지 않는 feature

('남여', '26')	0.006069	0.000673	0.009083	0.000957	0.001010	0.000583	0.007647	0.001334	0.002066	0.000568
('prod_size', '38')	0.004831	0.000366	0.004773	0.000263	0.006852	0.000472	0.020104	0.000899	0.010461	0.000279
('마더코드', '1')	0.003781	0.000892	0.002976	0.000612	0.009492	0.001901	0.025406	0.002381	0.018939	0.001380
('상품명', '3')	0.003761	0.001115	0.011189	0.002309	0.035423	0.002235	0.029453	0.001641	0.003510	0.000691
('상품코드', '2')	0.002723	0.000166	0.000143	0.000959	0.012984	0.000890	0.021071	0.000793	0.011140	0.000999
('season_prod', '28')	0.002634	0.000680	0.007911	0.000908	0.001700	0.000360	0.003912	0.000658	0.000054	0.000623
('상품군', '4')	0.002272	0.000178	0.001478	0.000099	-0.001144	0.001331	0.005600	0.000433	0.001190	0.001435
('브랜드', '24')	0.002060	0.000460	-0.002535	0.000573	0.021044	0.000530	0.008413	0.000987	-0.001741	0.000789
('취급액_mean_quantile_10', '29')	0.001982	0.000623	-0.000851	0.000438	0.005328	0.000890	0.005521	0.000753	0.005306	0.000403
('중분류', '23')	0.001630	0.000467	-0.000750	0.000258	0.008624	0.000877	0.003643	0.000801	0.003524	0.001121
('소분류', '22')	0.001474	0.000462	0.002725	0.000794	0.004187	0.001090	0.009343	0.000894	-0.000251	0.000624
('기온', '12')	0.001035	0.000184	0.000403	0.000280	-0.000071	0.000731	0.003141	0.000958	-0.000571	0.000777
('weekend', '36')	0.000637	0.000210	0.000737	0.000331	0.002614	0.000133	0.000126	0.000081	0.002605	0.000698
('뉴스시청률', '20')	0.000480	0.000135	0.002871	0.000568	0.002456	0.000539	0.003574	0.001096	-0.003340	0.000257
('시가지수', '19')	0.000135	0.000043	-0.000273	0.000114	-0.001037	0.000257	0.003189	0.000861	0.000686	0.000418
('요일', '6')	0.000099	0.000046	0.000000	0.000000	0.000412	0.000213	0.000560	0.000440	-0.000038	0.000175
('누적강수량', '13')	0.000043	0.000059	0.000242	0.000099	-0.001177	0.000920	0.003071	0.000488	-0.000072	0.000241
('등락률(%)', '17')	0.000023	0.000024	0.000204	0.000058	-0.000063	0.000170	0.000454	0.000149	-0.000126	0.000194
('expensive', '33')	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
('season', '10')	0.000000	0.000000	0.000000	0.000000	-0.000018	0.000260	0.000400	0.000231	-0.000074	0.000071
('middle_class', '37')	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
('cheap0', '35')	-0.000022	0.000007	0.000000	0.000000	0.000000	0.000000	0.000007	0.000087	0.000081	0.000035
('풍속', '14')	-0.000023	0.000117	0.000149	0.000084	-0.000452	0.000199	0.000839	0.000809	-0.000408	0.000112

취급액 예측 모델

Feature selection

: Permutation importance 결과 (상품군 별 중요한 feature)

	중요한 feature
전체	취급액_mean_quantile_10_소분류, group, muil, hour, holiday_yn, 노출(분), 판매단가, prod_size,
의류	취급액_mean_quantile_10_소분류, group, 노출(분), 판매단가, hour
속옷	취급액_mean_quantile_10_소분류, group, hour, 노출(분), holiday_yn
주방	Muil, group, 취급액_mean_quantile_10_소분류, prod_size, 취급액_mean_log_quantile_10_소분류, hour, holiday_yn, brand
농수축	Group, holiday_yn, hour, 취급액_mean_quantile_10_소분류, 소분류
이미용	Group, 취급액_mean_quantile_10_소분류, hour, holiday_yn, hour, 노출(분)
가전	Group, 취급액_mean_quantile_10_소분류, 취급액_mean_quantile_10, 취급액_mean_log_quantile_10_소분류, 판매단가, hour, 소분류
생활용품	Group, 취급액_mean_quantile_10_소분류, muil, , 취급액_mean_log_quantile_10_소분류, 판매단가, hour, holiday_yn, 기온
건강기능	Group, hour, 뉴스시청률, 판매단가, 취급액_mean_quantile_10_소분류, 취급액_mean_quantile_10,
잡화	취급액_mean_quantile_10_소분류, group, hour, 판매단가, holiday_yn, , 취급액_mean_log_quantile_10_소분류, muil, 노출(분)
가구	Group, 상품명, , 취급액_mean_quantile_10_소분류, 마더코드, brand, 판매단가, muil, prod_size
침구	취급액_mean_log_quantile_10_소분류, group, hour, 소분류, prime_time, 상품명,

취급액 예측 모델

	Model1	Model2	Model3
Dataset	Dataset selection <u>상품군 별 dataset 분리</u>	Dataset selection <u>상품군 별 dataset 분리 X</u>	시간 관련 변수, 상품명
방식	LGBM, XGB 앙상블	LGBM, XGB, Catboost 앙상블	TextCNN
성능	28.46361870386592	30.71434560313055	31.045695804263573



Weighted sum

Model1 X 0.6 + Model2 X 0.1 + Model3 X 0.3

= 27.3173

- 문제정의

- 방송 편성 최적화

- 취급액 예측 모델

- **방송 편성 최적화**

- 의의 및 발전 방향

방송 편성 최적화

편성방안 (하루 기준)

1) 2020년 6월에 판매된 모든 상품의 **취급액을 20분 단위로 예측**한다.

(노출 분 feature 를 이용하기 위해 20분 단위로 나누어 준 것임.)

2) 실제 편성은 1시간 단위로 이루어져 있는 경우가 대부분 이므로, **3개씩 그룹을 묶어 취급액을 더해준다.** (최종적으로 60분 단위로 취급액을 예측하는 형태)

3) **헝가리안 알고리즘**을 통해 취급액이 가장 커지는 시간별, 상품 조합을 찾는다.

방송 편성 최적화

편성방안 (하루 기준)

- 1) 2020년 6월에 판매된 모든 상품의 취급액을 20분 단위로 예측한다.
(노출 분 feature 를 이용하기 위해 20분 단위로 나누어 준 것임.)

20분 단위로 예측



	잭필드 남성 반 팔셔츠 4 종	쿠미두니 카 쿨 레 이시 란 슈웨이퍼 &팬티	바비리스 퍼펙트 볼륨스타 일러	램프룩 자동화전 냄비	벨레즈온 심리스 원피스 4 종 패키지	AAC 삼채 포기김치 10kg	아키 라 이크라 릴렉스 보정브라 패키지 (뉴아키 28차)	에이유플 러스 슈 퍼션스틱 1004(최 저가)	[기간] 제 주바다자 연산돔39 마리	한라궁 황칠 제 주오메기 떡 40개 (60g * 40봉)	한라궁 황칠 제 주오메기 떡 2종 40개 (호 박오메기 20봉+팔 오메기20 봉)	LG전자 통 돌이 세탁 기 TR14WK1(화 이트)	코몽트 남성 티 셔츠 8종 (시즌1)
2020-06-01 06:40:00	5141458	14762280	12409600	11942391	11406869	12595533	13216264	13301211	13516697	13210069	17052836	16652915	16242592
2020-06-01 07:00:00	19624512	19540230	22889829	19540600	19780024	17710185	11347738	13693368	13970029	11140795	11810285	8335976	23209759
2020-06-01 07:20:00	81996147	91846211	59476053	59971128	60785868	53201390	67745050	69930655	47847064	50755253	66752898	49169424	46992152
2020-06-01 07:40:00	38377028	28552095	35961017	34093247	32648497	36185681	32351451	32883203	31748757	30307685	32257728	27036806	27205206
2020-06-01 08:00:00	43477505	36374906	37859733	41122075	41021322	40558055	46525680	39466798	38485377	47877768	44346924	50171137	38191700

방송 편성 최적화

편성방안 (하루 기준)

- 2) 실제 편성은 1시간 단위로 이루어져 있는 경우가 대부분 이므로, 3개씩 그룹을 묶어
취급액을 더해준다. (최종적으로 60분 단위로 취급액을 예측하는 형태)

60분 단위로 예측



	잭필드 남 성 반팔셔 츠 4중	쿠미투니카 쿨 레이시 안쥬웨어 &팬티	바비리스 퍼펙트 볼 룸스타일 러	램프국 자 동회전냄비	벨레즈온 심리스 원 피스 4중 패키지	AAC 삼채포 기김치 10kg	아키 라이 크라 알렉 스 보정브 라 패키지 (뉴아키28 차)	에이유플러 스 슈퍼선 스틱 1004(최저 가)	[기간] 제주 바다자연산 돔39마리	한라궁 황 칠 제주오 메기떡 40 개 (60g * 40봉)	한라궁 황 칠 제주오 메기떡 2중 40개 (호박 오메기20봉 +팔오메기 20봉)	LG전자 통 돌이 세탁 기 TR14WK1(화 이트)	코몽트 남 성 티셔츠 8중 (시즌1)
2020-06-01 06:40:00	106762117.0	126148721.0	94775482.0	91454119.0	91972761.0	83507108.0	92309052.0	96925234.0	75333790.0	75106117.0	95616019.0	74158315.0	86444503.0
2020-06-01 07:40:00	109473986.0	91490938.0	98121505.0	102526585.0	102316483.0	111557306.0	112821389.0	108809879.0	110553821.0	113156933.0	111957041.0	116857588.0	96692910.0
2020-06-01 08:40:00	97224336.0	81338716.0	68824108.0	75696733.0	90162550.0	72343092.0	53152817.0	58003078.0	38648113.0	43259966.0	43998171.0	41037097.0	49627552.0
2020-06-01 09:40:00	32515612.0	37642676.0	44005356.0	38012105.0	46231489.0	51098959.0	46153280.0	42569252.0	47115315.0	50236169.0	53418300.0	51092082.0	55646803.0
2020-06-01 10:40:00	75979641.0	67869422.0	63471896.0	66225002.0	68821384.0	68851591.0	67585017.0	57006594.0	64830364.0	57615648.0	56087529.0	58350519.0	56318941.0
2020-06-01 11:40:00	40664607.0	40250500.0	35625496.0	31422726.0	31109462.0	37510529.0	33853974.0	35542928.0	51652983.0	46370683.0	53411336.0	51496892.0	49864914.0

방송 편성 최적화

편성방안 (하루 기준)

3) 헝가리안 알고리즘을 통해 시간별, 상품별 매출을 최적화하는 조합을 탐색

	product	weight
2020-06-01 06:40:00	에버라스트 에버 백팩 4종	170212407.0
2020-06-01 07:40:00	추억의쥐치포110장	132351100.0
2020-06-01 08:40:00	CERINI by PAT 남성 에어홀 팬츠3종	98412242.0
2020-06-01 09:40:00	비비고 만두	137653834.0
2020-06-01 10:40:00	무이자 삼성 UHD TV 65형 KU65UT7000FXKR	116301626.0
2020-06-01 11:40:00	무이자 20년 무풍 슬림 16형 화이트(절전) 홈멀티 (AF16T5774WZRT) ...	121474622.0
2020-06-01 12:40:00	일시불 am마카롱 보냉보온 텀블러세트	116912121.0
2020-06-01 13:40:00	기라로쉬 멀티 클립온 선글라스	62555958.0
2020-06-01 14:40:00	LSX 아쿠아엑스 란쥬 셰이퍼 패키지	76910913.0
2020-06-01 15:40:00	일시불 파격찬스 굿프렌드 스트레칭 안마 매트 GOOD-A7	125096219.0
2020-06-01 16:40:00	신일써큘레이터 스탠드 화이트+화이트(SIF-R09DWH)	106092183.0

편성표

참고

편성표 (하루기준)

3) 헝가리안 알고리즘을 통해 취급액이 가장 커지는 시간별, 상품 조합을 찾는다.

다음과 같은 행렬에서 열, 행이 중복되지 않게, 가장 큰 조합을 찾아주는 알고리즘

		product			weight
2020-06-01 06:40:00		에버라스트 에버 백팩 4종			170212407.0
2020-06-01 07:40:00					100.0
2020-06-01 08:40:00	i_1	j_1	j_2	j_3	142.0
2020-06-01 09:40:00	i_2	3	8	9	134.0
2020-06-01 10:40:00	i_3	4	12	7	126.0
2020-06-01 11:40:00		4	8	5	121474622.0
2020-06-01 12:40:00		무이자 20년 무풍 슬림 16형 화이트(절전) 홈멀티 (AF16T5774WZRT) ...			116912121.0
2020-06-01 13:40:00		일시불 am마카롱 보냉보온 텀블러세트			62555958.0
2020-06-01 14:40:00		기라로쉬 멀티 클립온 선글라스			76910913.0
2020-06-01 15:40:00		LSX 아쿠아엑스 란쥬 쉐이퍼 패키지			125096219.0
2020-06-01 16:40:00		일시불 파격찬스 굿프렌드 스트레칭 안마 매트 GOOD-A7			106092183.0
2020-06-01 17:40:00		신일써클레이터 스탠드 화이트+화이트(SIF-R09DWH)			

편성표

요일별 편성결과 분석

<주말>

<토요일>

2020-06-06 06:40:00	지나송 보노 화이트에디션 암막 커튼(중형)	166264133
2020-06-06 07:40:00	지나송 보노 화이트에디션 암막 커튼(슈퍼특대형)	142814722
2020-06-06 08:40:00	신모델 무풍 슬림 16형 화이트(절전) 스탠드(AF16T5774WZT) + 삼성 선풍기 2대 (SF	132221847
2020-06-06 09:40:00	지나송 로사 웨딩스타일 암막 이중커튼(중형)	153245761
2020-06-06 10:40:00	한스데코 샤를 이중 암막 레이스 커튼(중형)	94875677
2020-06-06 11:40:00	초특가 공간아트 매직건조대 대중소 종합 세트	195134076

<일요일>

2020-06-07 06:40:00	보루네오 델루나 유로탑 슬라이딩 LED침대 킹	137200533
2020-06-07 07:40:00	무이자 삼성 UHD TV 75형 / KU75UT7000FXKR	151232309
2020-06-07 08:40:00	푸마 드라미셀 썸머 넷서 드로스 11송 패키지	152737589
2020-06-07 09:40:00	와도 피마환전복 1.2kg	144945476
2020-06-07 10:40:00	신쿠첸압력밥솥 10인용	102367176
2020-06-07 11:40:00	초특가 공간아트 매직건조대 대중소 종합 세트	189267756

<평일>

<금요일>

2020-06-05 06:40:00	지나송 보노 화이트에디션 암막 커튼(중형)	139159086
2020-06-05 07:40:00	신일써클레이터 스탠드 블랙+블랙(SIF-PC30DCC)	127940987
2020-06-05 08:40:00	엘렌실라&코즈미 섹클렌저(펌핑기)	141642405
2020-06-05 09:40:00	QISS 프리미엄어 아침애플리다 마시는한끼 42통	124522196
2020-06-05 10:40:00	[골로올레] 제주 당도선별 하우스감귤 1.1kg 3박스	95732701
2020-06-05 11:40:00	마르엘라로사티 플로럴썸머원피스3종	144863958

<화요일>

2020-06-02 06:40:00	LG 울트라HD TV AI ThinQ(인공지능 씽큐) 75형 75UN7850KNA	122340784
2020-06-02 07:40:00	지나송 로사 웨딩스타일 암막 이중커튼(특대형)	144042992
2020-06-02 08:40:00	갯바위전장김27봉	167988779
2020-06-02 09:40:00	벤티리 폴딩 의자 1개	126683641
2020-06-02 10:40:00	무이자 매직쉐프 플렉타인 전자레인지	104900204
2020-06-02 11:40:00	마르엘라로사티 플로럴썸머원피스3종	162899520

→ 주말 오전: **가전, 가구, 침구**의 편성이 유리

→ 평일 오전: 주말 오전과는 다르게 **잡화, 농수축, 의류** 등의 상품군이 침구 등에 비해 상대적으로 많이 등장

방송 편성 최적화

시간대별 편성 결과 분석

- 16시~20시

<기준>

2019-06-07 16:00	20	100532	201608	" 100%리얼 착즙 석류 ", 총 4박스	건강기능	77,000	28,098,000
2019-06-07 16:20	20	100532	201608	" 100%리얼 착즙 석류 ", 총 4박스	건강기능	77,000	33,281,000
2019-06-07 16:40	20	100532	201608	" 100%리얼 착즙 석류 ", 총 4박스	건강기능	77,000	49,706,000
2019-06-07 17:00	20	100644	201964	사용불가미리구운 자연산 바다장어 9팩	농수축	49,900	33,642,000
2019-06-07 17:20	20	100644	201964	사용불가미리구운 자연산 바다장어 9팩	농수축	49,900	43,540,000
2019-06-07 17:40	20	100644	201964	사용불가미리구운 자연산 바다장어 9팩	농수축	49,900	45,858,000
2019-06-07 19:00	20	100715	202090	젠트웰 스트레치 통풍 메쉬 자켓 1종	의류	49,000	12,727,000
2019-06-07 19:20	20	100715	202090	젠트웰 스트레치 통풍 메쉬 자켓 1종	의류	49,000	29,234,000
2019-06-07 19:40	20	100715	202090	젠트웰 스트레치 통풍 메쉬 자켓 1종	의류	49,000	41,863,000
2019-06-07 20:00	20	100149	200539	일시불 LG 휘센 씽큐 에어컨 2IN1형 FQ17	가전	2,209,000	10,719,000
2019-06-07 20:00		100149	200471	무이자 LG 휘센 씽큐 에어컨 2IN1형 FQ17	가전	2,299,000	16,708,000
2019-06-07 20:00		100149	200542	일시불 LG 휘센 씽큐 에어컨 스탠드형FQ17	가전	1,899,000	
2019-06-07 20:00		100149	200474	무이자 LG 휘센 씽큐 에어컨 스탠드형FQ17	가전	1,989,000	

방송 편성 최적화

시간대별 편성 결과 분석

- 16시~20시

<기준>

2019-06-12 16:00	20	100644	201964	사용불가미리구운 자연산 바다장어 9팩	농수축	49,900	21,464,000
2019-06-12 16:20	20	100644	201964	사용불가미리구운 자연산 바다장어 9팩	농수축	49,900	34,471,000
2019-06-12 16:40	20	100644	201964	사용불가미리구운 자연산 바다장어 9팩	농수축	49,900	43,391,000
2019-06-12 17:00	20	100253	200864	안동간고등어 20팩	농수축	30,900	48,713,000
2019-06-12 17:20	20	100253	200864	안동간고등어 20팩	농수축	30,900	57,498,000
2019-06-12 17:40	20	100253	200864	안동간고등어 20팩	농수축	30,900	62,771,000
2019-06-12 18:00	20	100167	200589	유귀열의 귀한 갯김치3kg+총각김치3kg (중)	농수축	38,900	29,052,000
2019-06-12 18:20	20	100167	200589	유귀열의 귀한 갯김치3kg+총각김치3kg (중)	농수축	38,900	47,525,000
2019-06-12 18:40	20	100167	200589	유귀열의 귀한 갯김치3kg+총각김치3kg (중)	농수축	38,900	54,633,000
2019-06-12 19:00	20	100618	201915	헤드 파이로라이트 드로즈 10종 패키지	속옷	69,000	17,453,000
2019-06-12 19:00		100618	201918	헤드 파이로라이트 트렁크 8종 패키지	속옷	69,000	7,413,000
2019-06-12 19:20	20	100618	201915	헤드 파이로라이트 드로즈 10종 패키지	속옷	69,000	34,460,000
2019-06-12 19:20		100618	201918	헤드 파이로라이트 트렁크 8종 패키지	속옷	69,000	8,577,000
2019-06-12 19:40	20	100618	201915	헤드 파이로라이트 드로즈 10종 패키지	속옷	69,000	49,386,000
2019-06-12 19:40		100618	201918	헤드 파이로라이트 트렁크 8종 패키지	속옷	69,000	16,975,000
2019-06-12 20:00	30	100305	200990	오모떼 프런트후크 프레쉬 데오브라 시즌4	속옷	139,000	38,662,000
2019-06-12 20:30	30	100305	200990	오모떼 프런트후크 프레쉬 데오브라 시즌4	속옷	139,000	88,495,000

방송 편성 최적화

시간대별 편성 결과 분석

- 16시~20시

<기준>

먹거리와 관련된 **농수축**
상품이 저녁시간대에 많이
편성 되어있음.

2019-06-20 16:00	20	100511	201575	자연산 손질 통오징어 21미	농수축	40,900	23,636,000
2019-06-20 16:20	20	100511	201575	자연산 손질 통오징어 21미	농수축	40,900	41,101,000
2019-06-20 16:40	20	100511	201575	자연산 손질 통오징어 21미	농수축	40,900	43,512,000
2019-06-20 17:00	20	100199	200671	소들넉 소 갈비탕 10팩	농수축	50,900	25,493,000
2019-06-20 17:20	20	100199	200671	소들넉 소 갈비탕 10팩	농수축	50,900	40,545,000
2019-06-20 17:40	20	100199	200671	소들넉 소 갈비탕 10팩	농수축	50,900	42,272,000
2019-06-20 18:00	20	100253	200864	안동간고등어 20팩	농수축	30,900	22,083,000
2019-06-20 18:20	20	100253	200864	안동간고등어 20팩	농수축	30,900	29,995,000
2019-06-20 18:40	20	100253	200864	안동간고등어 20팩	농수축	30,900	40,810,000
2019-06-20 19:00	30	100683	202032	무이자 매직셰프 10리터 듀얼쿡 에어프라이	주방	164,000	6,569,000
2019-06-20 19:00		100683	202035	일시불 매직셰프 10리터 듀얼쿡 에어프라이	주방	154,000	5,048,000
2019-06-20 19:30	30	100683	202032	무이자 매직셰프 10리터 듀얼쿡 에어프라이	주방	164,000	22,504,000
2019-06-20 19:30		100683	202035	일시불 매직셰프 10리터 듀얼쿡 에어프라이	주방	154,000	28,909,000

방송 편성 최적화

시간대별 편성 결과 분석

- 16시~20시

<예측>

예측 결과 역시 **비슷한**

상품군을 편성해주는 모습

을 볼 수 있음.

A	B	C
	product	weight
2020-06-01 6:40:00	에버라스트 에버 백팩 4종	170212407
2020-06-01 7:40:00	추억의퀴치포110장	132351100
2020-06-01 8:40:00	CERINI by PAT 남성 에어홀 팬츠3종	98412242
2020-06-01 9:40:00	비비고 만두	137653834
2020-06-01 10:40:00	무이자 삼성 UHDTV 65형 KU65UT7000FXKR	116301626
2020-06-01 11:40:00	무이자 20년 무풍 슬림 16형 화이트(절전) 홈멀티 (AF16T5774WZRT) + 삼성 선풍기 2대(SFN-M35GABL)	121474622
2020-06-01 12:40:00	일시불 am마카롱 보냉보온 텀블러세트	116912121
2020-06-01 13:40:00	기라로쉬 멀티 클립온 선글라스	62555958
2020-06-01 14:40:00	LSX 아쿠아엑스 란쥬 쉐이퍼 패키지	76910913
2020-06-01 15:40:00	일시불 파격찬스 굿프렌드 스트레칭 안마 매트 GOOD-A7	125096219
2020-06-01 16:40:00	산일씨클레이터 스탠드 화이트+화이트(SIF-R09DWH)	106092183
2020-06-01 17:40:00	린나이 넷지 가스레인지 2구 (압력솥)	123397877
2020-06-01 18:40:00	무이자 20년 무풍 슬림 16형 화이트(절전) 스탠드(AF16T5774WZT) + 삼성 선풍기 2대(SFN-M35GABL)	104347150
2020-06-01 19:40:00	완도특대참전목 16~18미	128476971
2020-06-01 20:40:00	일시불 20년 무풍 슬림 18형 화이트(절전) 스탠드(AF18T5774WZT) + 삼성 선풍기 2대(SFN-M35GABL)	125765027
2020-06-01 21:40:00	현대어찬 손질갈치 14팩	132645367
2020-06-01 22:40:00	프리미엄 클린샤워	94227530
2020-06-01 23:40:00	제오닉 나누담 밀폐용기	178588378
2020-06-02 0:40:00	(권) 행복돌침대 로젠 황토불침대	110101063
2020-06-02 1:40:00	맛의명가 반야월 튀김기 2kg + 돈볼로스 1kg	127265451

→ 주방
→ 가전
→ 농수축

방송 편성 최적화

시간대별 편성 결과 분석

- 10시~12시

<기존>

2019-06-07 10:00	20	100188	200624	코치 노리타 리스틀릿	잡화	138,000	4,368,000
2019-06-07 10:00		100188	200640	코치 아웃라인 파일백	잡화	138,000	2,188,000
2019-06-07 10:00		100188	200634	코치 시그니처 숄더백	잡화	148,000	2,730,000
2019-06-07 10:00		100188	200636	코치 시그니처 켈시 사첼	잡화	238,000	6,288,000
2019-06-07 10:20	20	100188	200624	코치 노리타 리스틀릿	잡화	138,000	12,655,000
2019-06-07 10:20		100188	200640	코치 아웃라인 파일백	잡화	138,000	9,427,000
2019-06-07 10:20		100188	200634	코치 시그니처 숄더백	잡화	148,000	11,696,000
2019-06-07 10:20		100188	200636	코치 시그니처 켈시 사첼	잡화	238,000	18,693,000
2019-06-07 10:40	20	100188	200624	코치 노리타 리스틀릿	잡화	138,000	20,203,000
2019-06-07 10:40		100188	200640	코치 아웃라인 파일백	잡화	138,000	9,424,000
2019-06-07 10:40		100188	200634	코치 시그니처 숄더백	잡화	148,000	13,427,000
2019-06-07 10:40		100188	200636	코치 시그니처 켈시 사첼	잡화	238,000	28,752,000
2019-06-07 11:00	20	100770	202250	무이자 예지리 스텐큐브 프라이팬	주방	179,000	11,918,000
2019-06-07 11:00		100770	202255	일시불 예지리 스텐큐브 프라이팬	주방	171,000	8,313,000
2019-06-07 11:20	20	100770	202250	무이자 예지리 스텐큐브 프라이팬	주방	179,000	36,484,000
2019-06-07 11:20		100770	202255	일시불 예지리 스텐큐브 프라이팬	주방	171,000	21,848,000
2019-06-07 11:40	20	100770	202250	무이자 예지리 스텐큐브 프라이팬	주방	179,000	36,977,000
2019-06-07 11:40		100770	202255	일시불 예지리 스텐큐브 프라이팬	주방	171,000	24,495,000
2019-06-07 12:00	20	100677	202023	2019 S/S 로베르타 디 까메리노 선글라스	잡화	69,800	23,176,000
2019-06-07 12:20	20	100677	202023	2019 S/S 로베르타 디 까메리노 선글라스	잡화	69,800	25,131,000
2019-06-07 12:40	20	100677	202023	2019 S/S 로베르타 디 까메리노 선글라스	잡화	69,800	38,025,000

방송 편성 최적화

시간대별 편성 결과 분석

- 10시~12시

<기준>

2019-06-12 10:00	20	100779	202276	쿠미투니카 라이크라 보정 레이스멀티팬티	속옷	39,900	15,395,000
2019-06-12 10:20	20	100779	202276	쿠미투니카 라이크라 보정 레이스멀티팬티	속옷	39,900	37,719,000
2019-06-12 10:40	20	100779	202276	쿠미투니카 라이크라 보정 레이스멀티팬티	속옷	39,900	50,723,000
2019-06-12 11:00	20	100139	200374	보몽드 샬비어 시어셔커 리플 침구세트 SS	침구	49,900	7,784,000
2019-06-12 11:00		100139	200373	보몽드 샬비어 시어셔커 리플 침구세트 C	침구	59,900	10,545,000
2019-06-12 11:00		100139	200372	보몽드 샬비어 시어셔커 리플 침구세트 K	침구	69,900	16,089,000
2019-06-12 11:20	20	100139	200374	보몽드 샬비어 시어셔커 리플 침구세트 SS	침구	49,900	12,945,000
2019-06-12 11:20		100139	200373	보몽드 샬비어 시어셔커 리플 침구세트 C	침구	59,900	24,499,000
2019-06-12 11:20		100139	200372	보몽드 샬비어 시어셔커 리플 침구세트 K	침구	69,900	19,313,000
2019-06-12 11:40	20	100139	200374	보몽드 샬비어 시어셔커 리플 침구세트 SS	침구	49,900	15,270,000
2019-06-12 11:40		100139	200373	보몽드 샬비어 시어셔커 리플 침구세트 C	침구	59,900	19,223,000
2019-06-12 11:40		100139	200372	보몽드 샬비어 시어셔커 리플 침구세트 K	침구	69,900	20,930,000
2019-06-12 12:00	20	100838	202461	무이자 쿠키 블랙스톤 전기밥솥 10인용	주방	278,000	6,230,000
2019-06-12 12:00		100838	202474	일시불 쿠키 블랙스톤 전기밥솥 10인용	주방	268,000	2,038,000
2019-06-12 12:00		100838	202462	무이자 쿠키 블랙스톤 전기밥솥 6인용	주방	248,000	6,487,000
2019-06-12 12:00		100838	202475	일시불 쿠키 블랙스톤 전기밥솥 6인용	주방	238,000	4,461,000
2019-06-12 12:20	20	100838	202461	무이자 쿠키 블랙스톤 전기밥솥 10인용	주방	278,000	16,532,000
2019-06-12 12:20		100838	202474	일시불 쿠키 블랙스톤 전기밥솥 10인용	주방	268,000	5,015,000
2019-06-12 12:20		100838	202462	무이자 쿠키 블랙스톤 전기밥솥 6인용	주방	248,000	2,804,000
2019-06-12 12:20		100838	202475	일시불 쿠키 블랙스톤 전기밥솥 6인용	주방	238,000	5,338,000
2019-06-12 12:40	20	100838	202461	무이자 쿠키 블랙스톤 전기밥솥 10인용	주방	278,000	20,650,000
2019-06-12 12:40		100838	202474	일시불 쿠키 블랙스톤 전기밥솥 10인용	주방	268,000	9,984,000
2019-06-12 12:40		100838	202462	무이자 쿠키 블랙스톤 전기밥솥 6인용	주방	248,000	12,905,000
2019-06-12 12:40		100838	202475	일시불 쿠키 블랙스톤 전기밥솥 6인용	주방	238,000	11,511,000

방송 편성 최적화

시간대별 편성 결과 분석

- 10시~12시

<기준>

주방,잡화, 침구가
많이 편성 되어있음.

2019-06-20 10:00	20	100097	200269	일시불 올리고 가스와이드그릴 프리미엄형	주방	119,000	8,718,000
2019-06-20 10:00		100097	200265	무이자 올리고 가스와이드그릴레인지 프리	주방	129,000	2,775,000
2019-06-20 10:20	20	100097	200269	일시불 올리고 가스와이드그릴 프리미엄형	주방	119,000	28,477,000
2019-06-20 10:20		100097	200265	무이자 올리고 가스와이드그릴레인지 프리	주방	129,000	9,507,000
2019-06-20 10:40	20	100097	200269	일시불 올리고 가스와이드그릴 프리미엄형	주방	119,000	51,371,000
2019-06-20 10:40		100097	200265	무이자 올리고 가스와이드그릴레인지 프리	주방	129,000	14,798,000
2019-06-20 11:00	20	100139	200374	보몽드 샬비어 시어서커 리플 침구세트 SS	침구	49,900	7,005,000
2019-06-20 11:00		100139	200373	보몽드 샬비어 시어서커 리플 침구세트 Q	침구	59,900	10,325,000
2019-06-20 11:00		100139	200372	보몽드 샬비어 시어서커 리플 침구세트 K	침구	69,900	12,908,000
2019-06-20 11:20	20	100139	200374	보몽드 샬비어 시어서커 리플 침구세트 SS	침구	49,900	10,163,000
2019-06-20 11:20		100139	200373	보몽드 샬비어 시어서커 리플 침구세트 Q	침구	59,900	18,850,000
2019-06-20 11:20		100139	200372	보몽드 샬비어 시어서커 리플 침구세트 K	침구	69,900	18,036,000
2019-06-20 11:40	20	100139	200374	보몽드 샬비어 시어서커 리플 침구세트 SS	침구	49,900	11,402,000
2019-06-20 11:40		100139	200373	보몽드 샬비어 시어서커 리플 침구세트 Q	침구	59,900	18,273,000
2019-06-20 11:40		100139	200372	보몽드 샬비어 시어서커 리플 침구세트 K	침구	69,900	22,466,000
2019-06-20 12:00	20	100158	200563	시스마르스 보헤미안 소가죽 샌들+물 슬러	잡화	49,800	19,338,000
2019-06-20 12:20	20	100158	200563	시스마르스 보헤미안 소가죽 샌들+물 슬러	잡화	49,800	28,182,000
2019-06-20 12:40	20	100158	200563	시스마르스 보헤미안 소가죽 샌들+물 슬러	잡화	49,800	43,373,000

방송 편성 최적화

시간대별 편성 결과 분석

- 10시~12시

<예측>

침구와 생활용품이

매출을 최적화하는 편성 상품군
으로 제시됨

2020-06-06 06:40:00	지나송 보노 화이트에디션 암막 커튼(중형)	166264133	
2020-06-06 07:40:00	지나송 보노 화이트에디션 암막 커튼(슈퍼특대형)	142814722	
2020-06-06 08:40:00	신모델 무풍 슬림 16형 화이트(절전) 스탠드(AF16T5774WZT) + 삼성 선풍기 2대 (SFN-M35GFGB)	132221847	
2020-06-06 09:40:00	지나송 로사 웨딩스타일 암막 이중커튼(중형)	153245761	침구
2020-06-06 10:40:00	한스데코 샤를 이중 암막 레이스 커튼(중형)	94875677	
2020-06-06 11:40:00	초특가 공간아트 매직건조대 대중소 종합 세트	195134076	생활용품
2020-06-06 12:40:00	마르엘라로사티 린넨 베스트 세트[3월]	78267122	
2020-06-06 13:40:00	에레키반1300 자석패치 6팩(총360매)	90549931	
2020-06-06 14:40:00	신모델 무풍 슬림 16형 화이트(절전) 홈멀티 (AF16T5774WZRT) + 삼성 선풍기 2대(SFN-M35GFGB)	139708623	
2020-06-06 15:40:00	제스프리 골드키워 48과	146075893	
2020-06-06 16:40:00	영광법성포굴비35마리	116902775	
2020-06-06 17:40:00	新쿠첸압력밥솥 6인용	102640551	
2020-06-06 18:40:00	일시불 am마카롱 보냉보온 텀블러세트	125164163	
2020-06-06 19:40:00	푸마 드라이셀 스트레치 트렁크 8중 패키지	104900411	
2020-06-06 20:40:00	AAC녹두삼계탕죽세트	120840161	
2020-06-06 21:40:00	무이자 LG전자 매직스페이스 냉장고	162929183	
2020-06-06 22:40:00	클라쎄 벽걸이 에어컨 TDOZ-S10JK	78553457	
2020-06-06 23:40:00	무이자 중근당건강 락토피트 생유산균골드 단하루	214047181	
2020-06-07 00:40:00	[루이띠에] 24K 999 순금 11.25g 더블 링크 체인 팔찌	141444122	
2020-06-07 01:40:00	(일)[보루네오] 피올레 천연소가죽 소파 3인용	125474124	

방송 편성 최적화

시간대별 편성 결과 분석

- 10시~12시

<예측>

오전에 농수축 판매 또한

예측 기반 매출 최적화 기반

편성 전략 중 하나

2020-06-03 09:40:00	완도꼬마활전복 1.3kg	132355572
2020-06-03 10:40:00	무이자 매직쉐프 플랫타입 전자레인지	99419836
2020-06-03 11:40:00	초특가 공간아트 매직건조대 대중소 종합 세트	150299214
2020-06-03 12:40:00	캐리어 벽걸이 에어컨 7형 ARC07VBC	59802793
2020-06-03 13:40:00	에레키반1300 자석패치 6팩(총360매)	78428192
2020-06-05 09:40:00	QISS 프리미어 아침애플리다 마시는한끼 42통	124522196
2020-06-05 10:40:00	[굴로올레] 제주 당도선별 하우스감굴 1.1kg 3박스	95732701
2020-06-07 09:40:00	완도꼬마활전복 1.3kg	144945476
2020-06-07 10:40:00	新쿠첸압력밥솥 10인용	102367176

방송 편성 최적화

상품별 편성 결과 분석

- 남성 셔츠, 속옷

<예측>

남성 셔츠, 속옷의 경우 저녁 및

아침 시간대에 해당 상품 매출

이 최적화되는 것으로 예측됨

	잭필드 남성 반팔셔츠 4종	쿠미투니카 쿨 레이시 란쥬웨어&팬티
2020-06-07 21:40:00	147998194	132132887
2020-06-07 23:40:00	108122358	110925977
2020-06-07 07:40:00	86646385	81459288
2020-06-07 09:40:00	78189282	64908740
2020-06-08 00:40:00	77535993	94864108
2020-06-07 20:40:00	73436470	69131496
2020-06-07 06:40:00	68836924	81851956
2020-06-08 01:40:00	67254013	65609166
2020-06-07 10:40:00	66607044	63431887
2020-06-07 14:40:00	64474523	72655292
2020-06-07 15:40:00	64237965	71650717
2020-06-07 19:40:00	55023968	53162233
2020-06-07 16:40:00	52387211	41073705
2020-06-07 12:40:00	49062512	49923773
2020-06-07 11:40:00	46730898	54772588
2020-06-07 08:40:00	44626408	49550133
2020-06-07 18:40:00	39522038	36548435
2020-06-07 17:40:00	36270852	37130087
2020-06-07 22:40:00	32003877	36262849
2020-06-07 13:40:00	26063608	36415318



- 문제정의
- 방송 편성 최적화
- 취급액 예측 모델
- 방송 편성 최적화

● 의의 및 발전 방향

의의 및 발전방향

[의의]

- 시간대, 상품 정보, 날씨, 시청률 등의 다양한 데이터 간의 관계를 파악하여 이를 기반으로 취급액을 예측
- 이를 바탕으로 기존의 전문가에 의존하는 정성적인 편성 방안이 아닌, 예상 취급액을 바탕으로 정량적인 방송 편성 방안 제시

[발전방향]

- 기존 데이터에 기반하고 있으므로, 아예 새로운 상품이 들어왔을 때의 대처가 어려움
→ 학습 데이터의 증대 및 전문가의 의견 등 정성적인 요소와의 결합으로 보완 가능
- 편성 최적화에 사용되는 데이터셋 구성 다양화

의의 및 발전방향

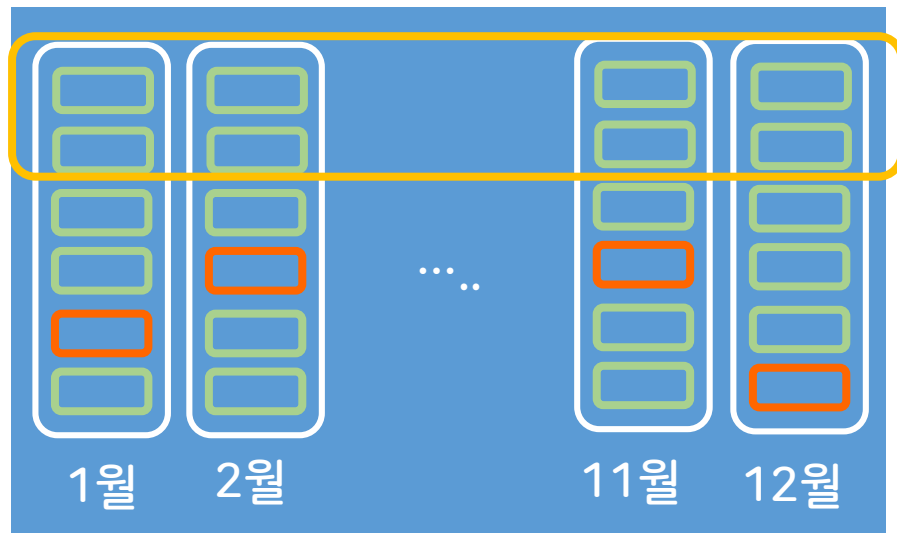
cf.) 편성 최적화 발전방향

해당 달의 잘 팔리는 상품 set,

한 해 동안 잘 팔린 상품 집합의 합집합으로 그 달의 편성 데이터셋을 구성함.

→ 편성데이터 셋 : ( +  + 기타)

→ 조금 더 전체 매출의 최적화하는 방안으로 기대됨.



전체 Dataset

항상 잘 팔리는 상품 : 전년도 기준 취급액이 항상 상위 30%에 드는 경우

월별로 잘 팔리는 상품 : 전년도 기준 월별 취급액이 높은 순 (상위 70%)



Thank you