

2020 빅 콘테스트 데이터분석 분야 챔피언리그

상품군별 모델 구축을 통한 NS Shop+ 취급액 예측



- 팀명 : 완판소년단
- 팀장 : 장영진 (jiuoo3@naver.com)
- 팀원 : 김민욱 (kmiiiaa@naver.com)
김재영 (livermore2025@gmail.com)
원종관 (jongkwan1@naver.com)
채종혁 (dnf15345@naver.com)

2020 빅 콘테스트 데이터분석 분야 챔피언리그

CONTENTS

01. 분석 배경 및 필요성

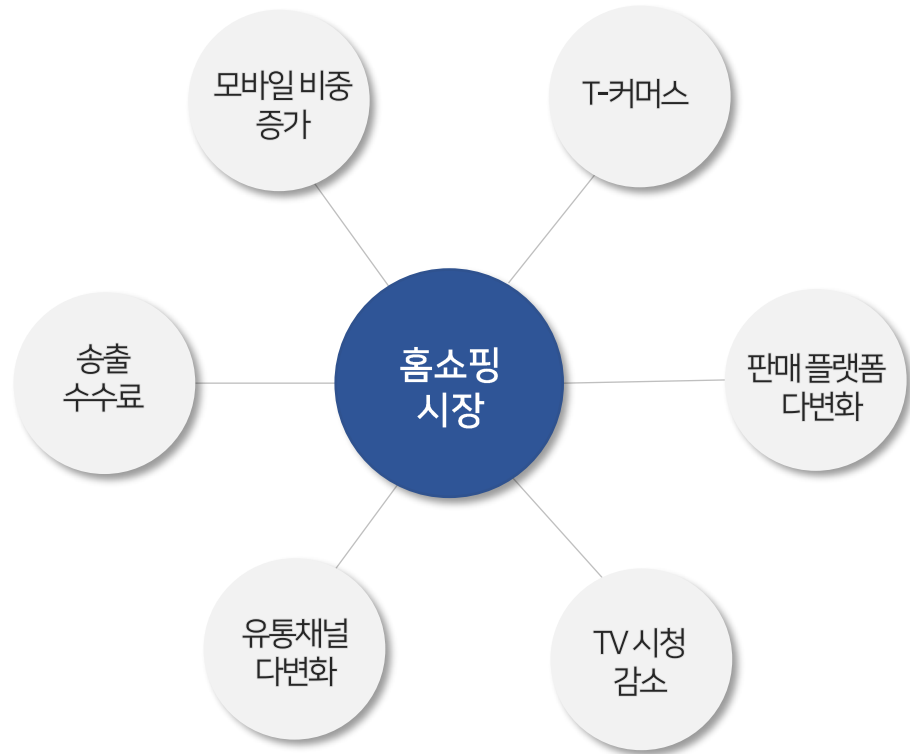
02. EDA

03. 데이터 전처리

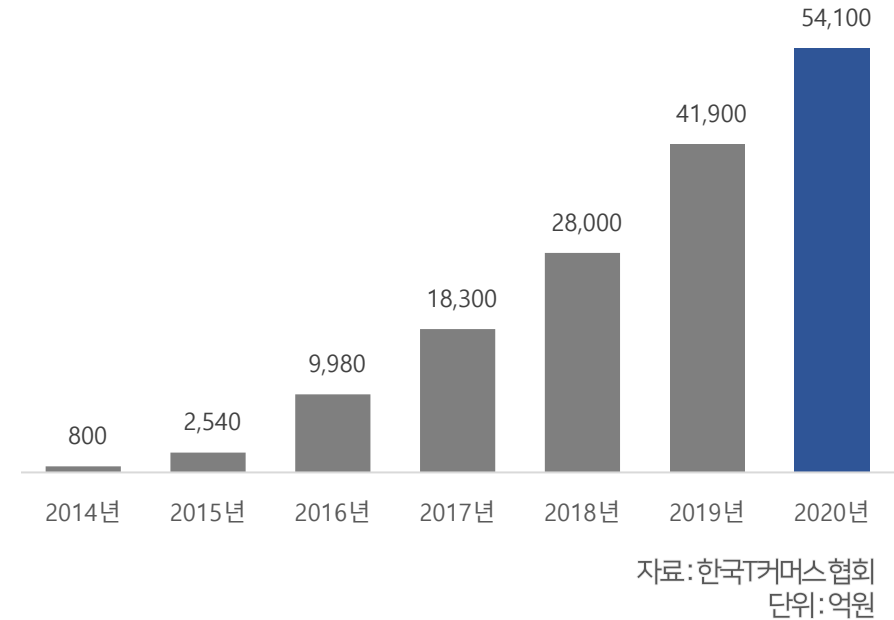
04. 모델 학습 및 파라미터 튜닝

05. 편성표 최적화 방안

● 홈쇼핑 시장 핵심 키워드



● T-커머스 연도별 취급액 변화



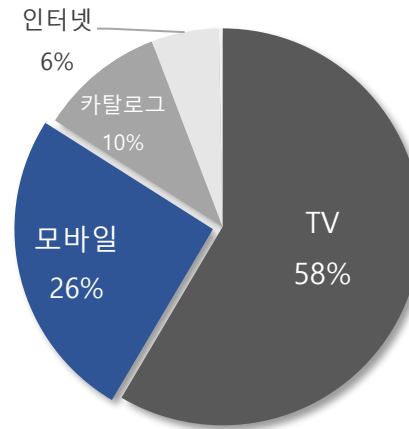
- 최근 홈쇼핑 업계는 코로나19 상황과 맞물려 급격한 내외 환경 변화를 마주함
- 각종 IT 기술이 적극 도입되면서 'T-커머스', '데이터 홈쇼핑', '라이브 커머스' 등 새로운 형태의 홈쇼핑이 활성화됨
- T-커머스는 빠르게 성장하고 있으며, 2020년 연 취급액이 5조 4천억원에 달할 것으로 예상됨

● NS 홈쇼핑 전년 대비 취급액



- 농수축산물 의무 편성 비율 60% → 식품 분야에 특화되어 있음
- 코로나 19의 영향으로 신선식품과 같은 주요 판매품목 인기
- 건강에 대한 소비자들의 관심 증가하면서 건강식품 구매 증가

● NS 홈쇼핑 쇼핑 수단 비중



- T-커머스, 라이브 커머스 활성화로 모바일 구매 비중 증가
- 언택트 소비 활성화로 모바일 사용량 증가
- 2030 세대의 홈쇼핑 이용량 증가

NS NS홈쇼핑

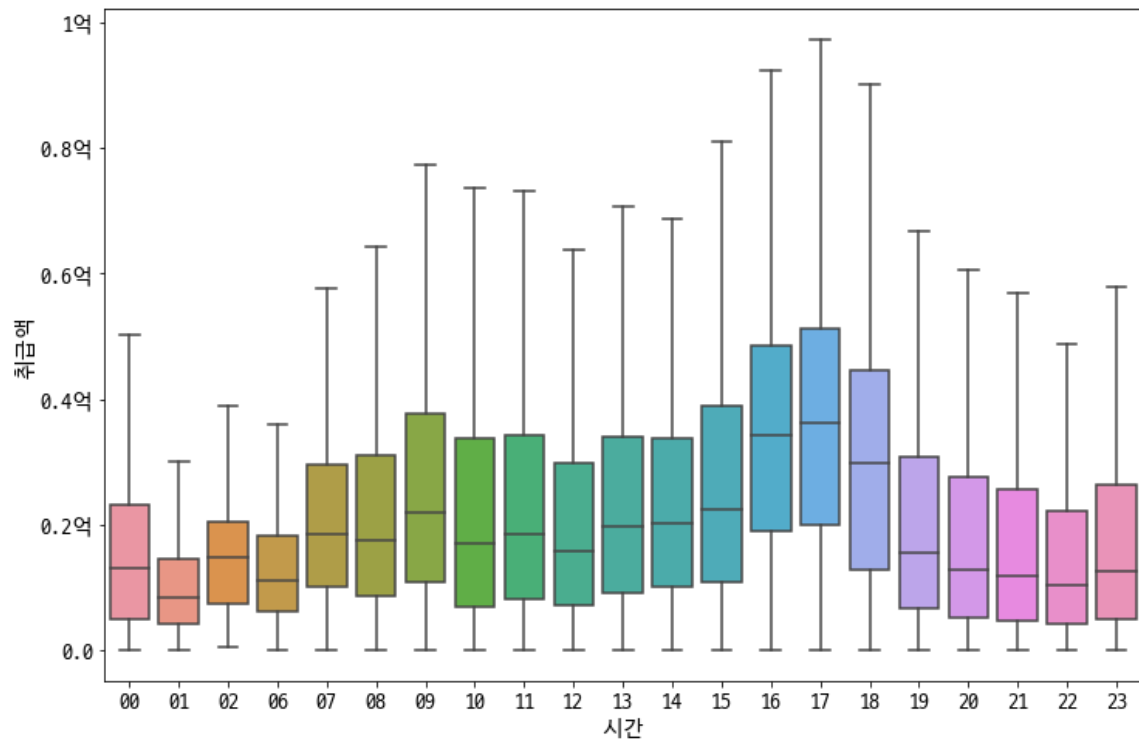
✓ 언택트 트렌드에 적합한 신규 고객 유치 & 기존 고객 이탈방지 전략 필요

NS Shop+

✓ 소비자들의 홈쇼핑 이용매체와 연령대가 다양해지고 있으므로, 다양한 소비자 특성을 반영해야 함

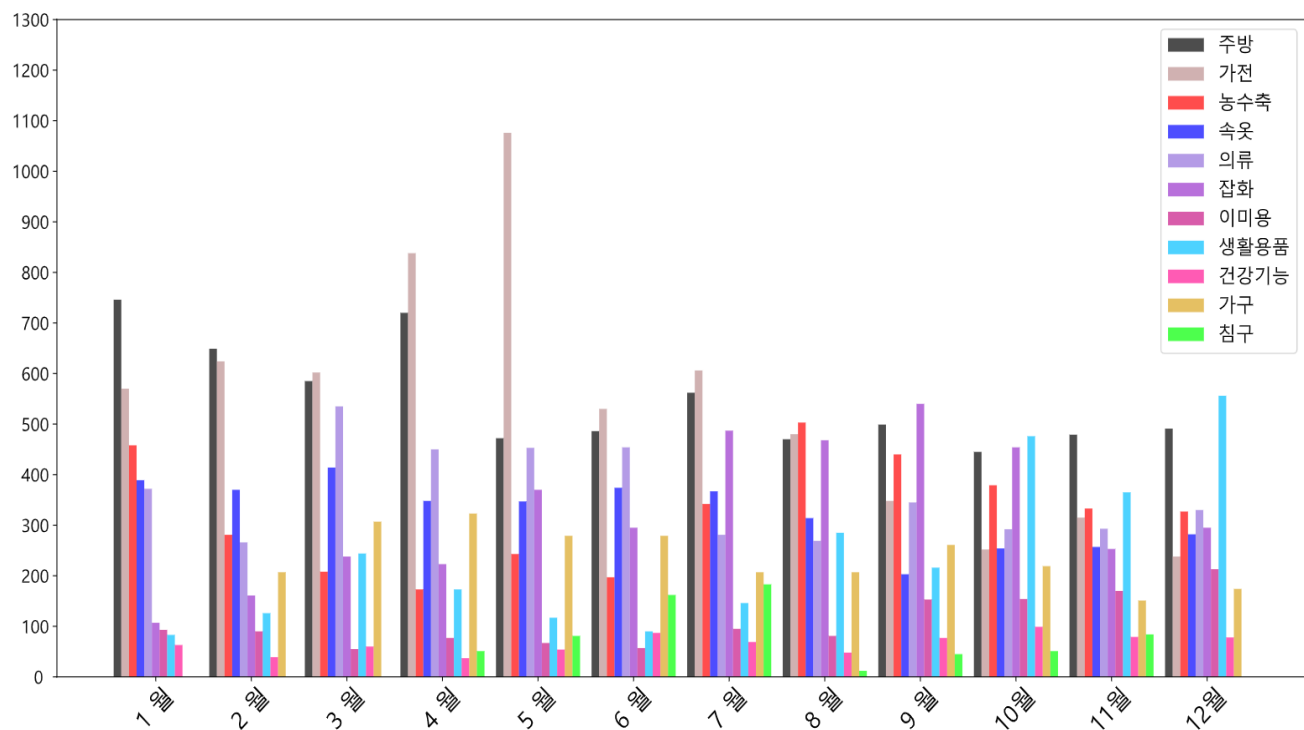
✓ 데이터 홈쇼핑은 녹화된 방송을 송출함 → 구매가능성이 높은 고객이 시청하도록 방송을 편성하는 것이 중요함

● 시간대별 취급액 합계 분포



- 주로 정오부터 늦은 오후(13시 ~ 18시) 까지 취급액이 많음
 - 새벽 시간대와 아침시간대에는 취급액 합계가 상대적으로 적음
- 시간대에 따라 상품의 취급액에 차이가 있음

● 월별 상품군에 따른 상품 판매 횟수

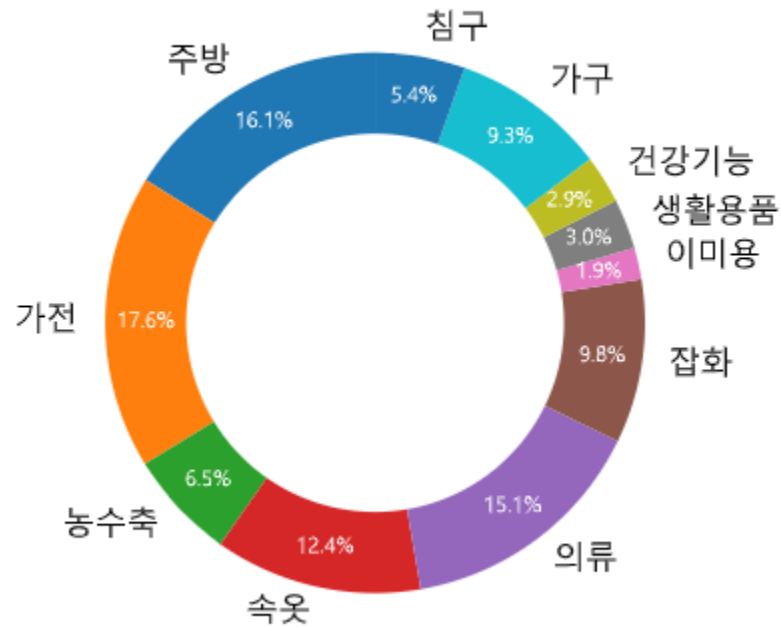


- 월별로 상품군마다 판매 횟수에서 상이한 패턴을 보이고 있음

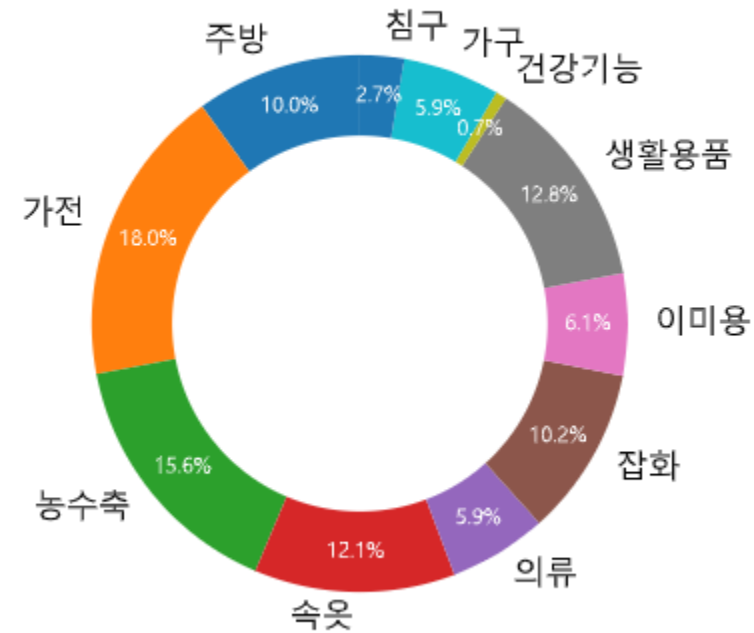
→ 상품군별로 계절성, 특정 이벤트 등의 영향 정도가 다름

- 2019년 6월 vs 2020년 6월의 상품군 분포 변화

The number of items by type, 2019-06

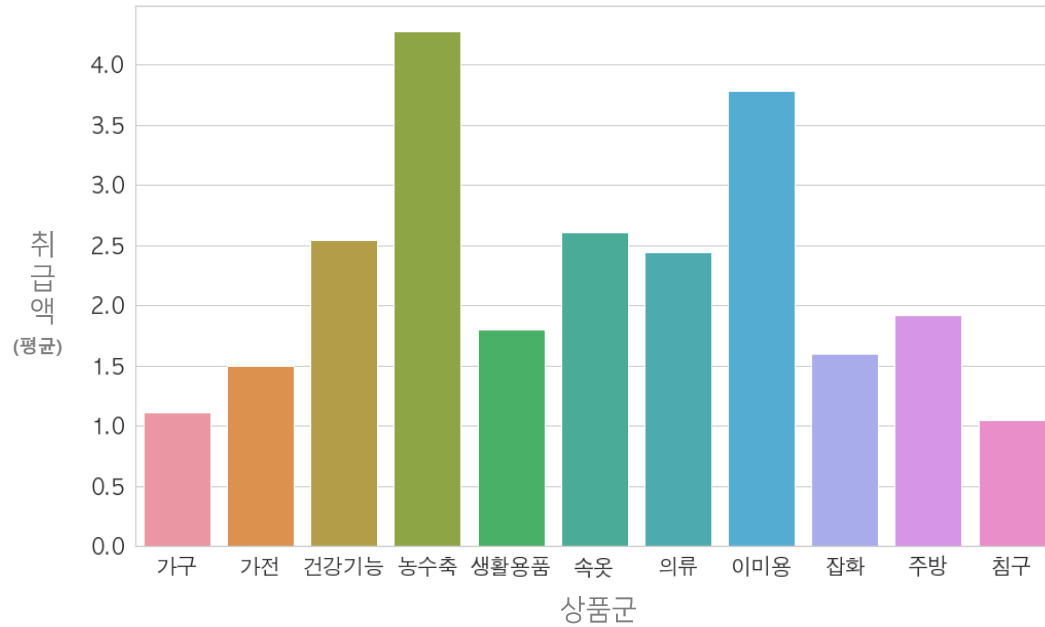


The number of items by type, 2020-06



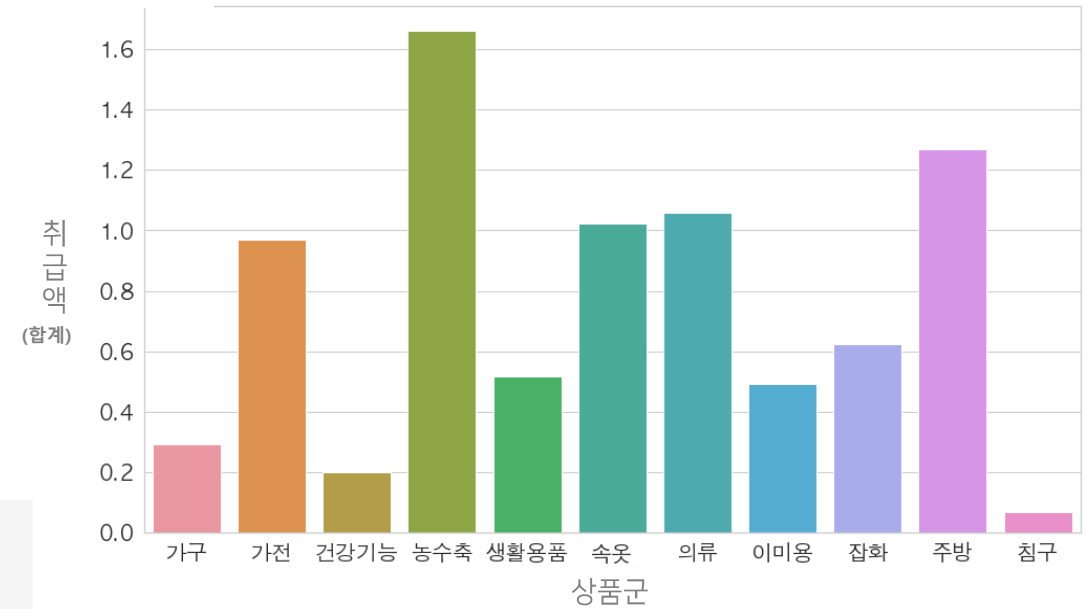
- 농수축, 생활용품, 이미용 상품군의 경우 판매 비중 증가함
- 주방, 가구, 의류, 건강기능, 침구 상품군의 판매 비중 감소함
- 가전, 잡화, 속옷 상품군의 경우 판매비율이 유사함

● 상품군별 취급액 **평균**



- 농수축, 이미용 카테고리의 평균 취급액이 높음
- 가구, 침구 상품군의 평균 취급액은 상대적으로 낮음

● 상품군별 취급액 **합계**



- 농수축, 주방, 가전 카테고리의 취급액 합계가 높음
- 건강기능, 침구 상품군의 취급액 합계가 낮음

✓ 상품군별로 취급액 평균과 합계에서 차이를 보임

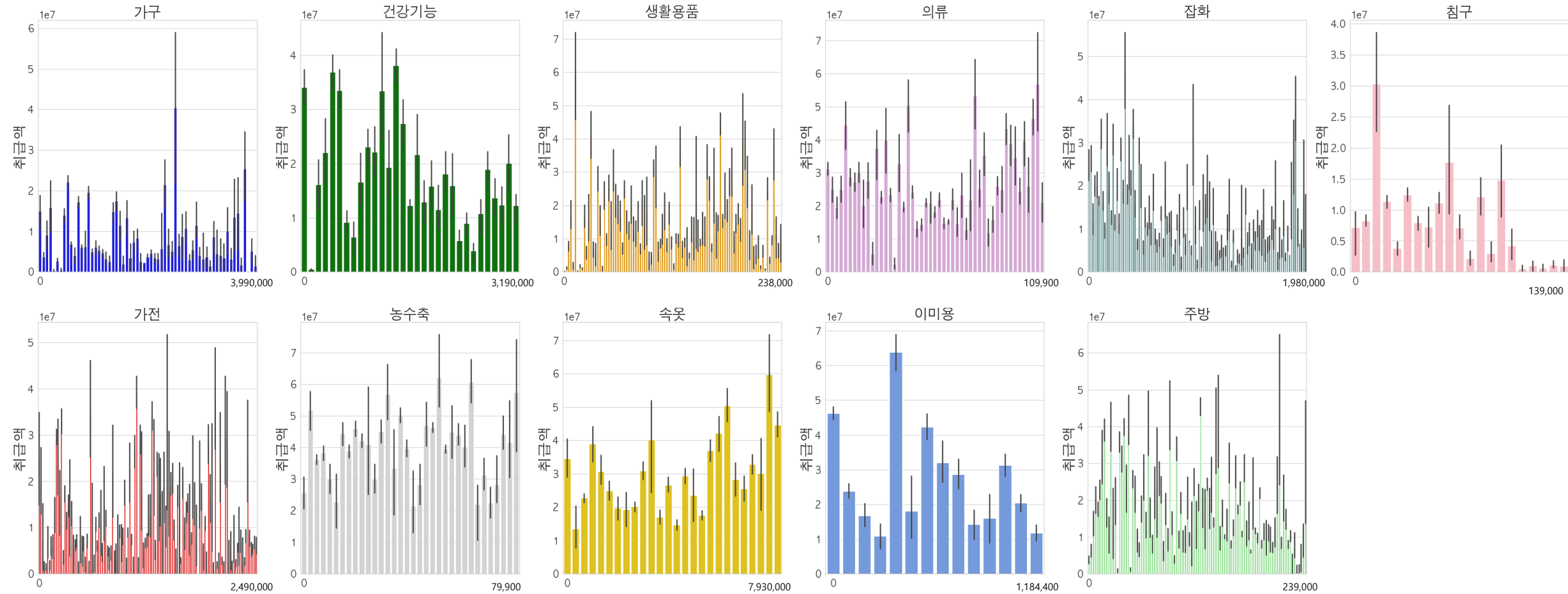
● 상품군별 취급액 분포 (x: 취급액, y: 편성횟수)



✓ 상품군별로 상이한 취급액이 분포를 보임

✓ 농수축, 이미용 상품군의 경우 상대적으로 취급액이 높은 방송이 많음

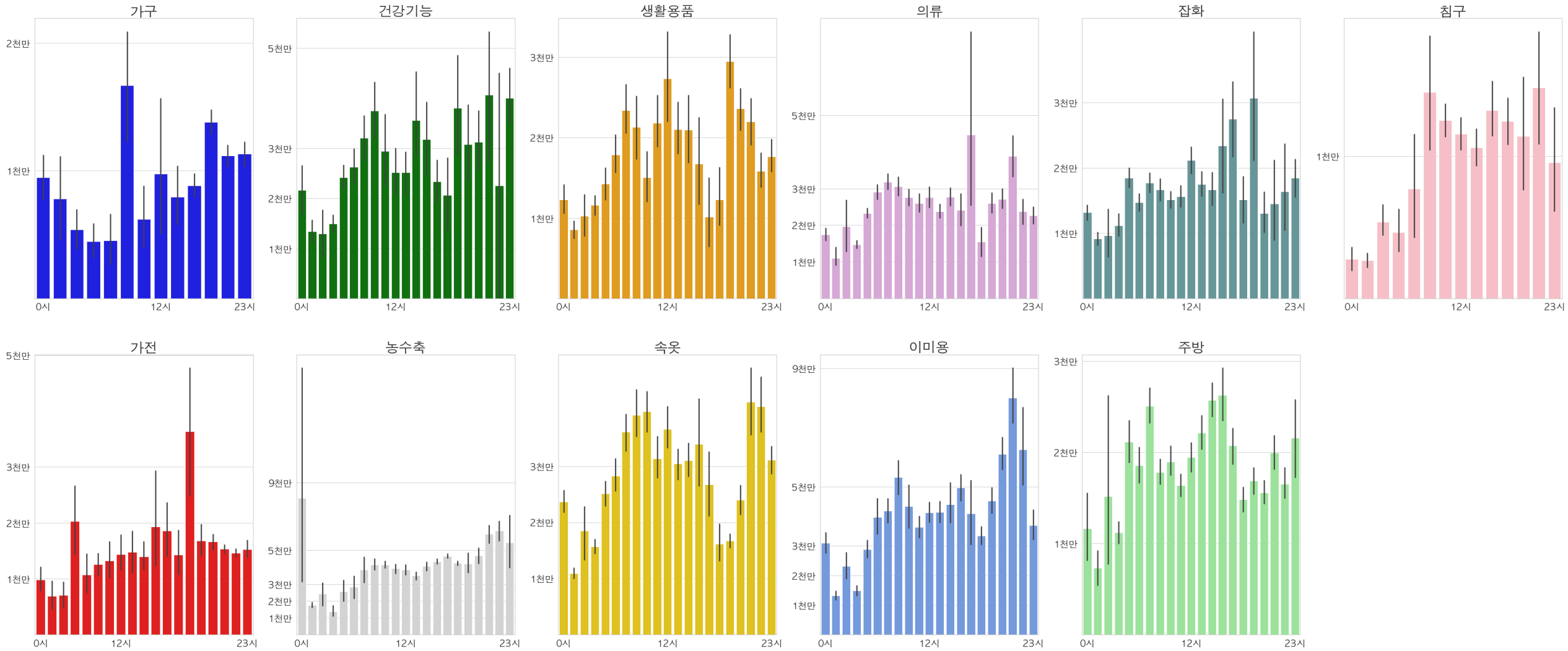
● 상품군별 판매단가에 따른 취급액 분포 (x: 판매단가, y: 취급액)



✓ 상품군별로 상품의 특징에 따라 상품 가격대가 다름

✓ 상품군별로 판매단가에 따른 취급액 분포를 살펴본 결과, 각각 다른 패턴을 보임

● 상품군별 시간대에 따른 취급액 분포 (x: 시간대, y: 취급액)



✓ 상품군별로 시간대에 따른 취급액 분포가 상이하므로 상품 특성에 따라 잘 팔리는 시간대가 다르다는 것을 알 수 있음

[EDA 결과]

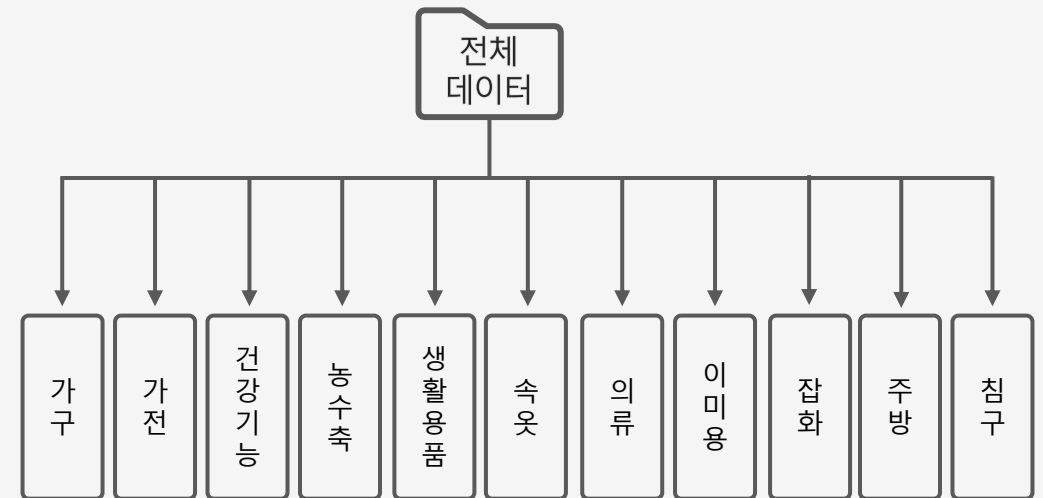
- 상품군에 따라서 판매단가, 주요 판매시간대, 계절성 등의 특징이 다름
→ 상품군별 특징, 취급액 패턴을 반영한 예측 필요

[상품 유형에 따른 구매행태 차이에 관한 연구]

- 소비자들은 소매업체와 상품특성에 따라 태도와 구매의도에서 차이를 보인다
- 온라인 쇼핑물과 오프라인 매장에서의 소비자의 태도차이는 '쇼핑가치'에 따라서 유의미한 차이를 보인다 (김원겸, 2004)
- 소비자들은 상품의 유형에 따라 정보 탐색 어려움 정도에서 차이를 보이며, 이는 구매 결정에 유의미한 영향을 미친다. (구명진 외 1명, 2015)
- 여러 상품 유형에 대하여 소비자들은 성별에 따라서 상품에 대한 관심도와 내재적 혁신성에서 유의미한 차이를 나타낸다. (정인희, 2015)

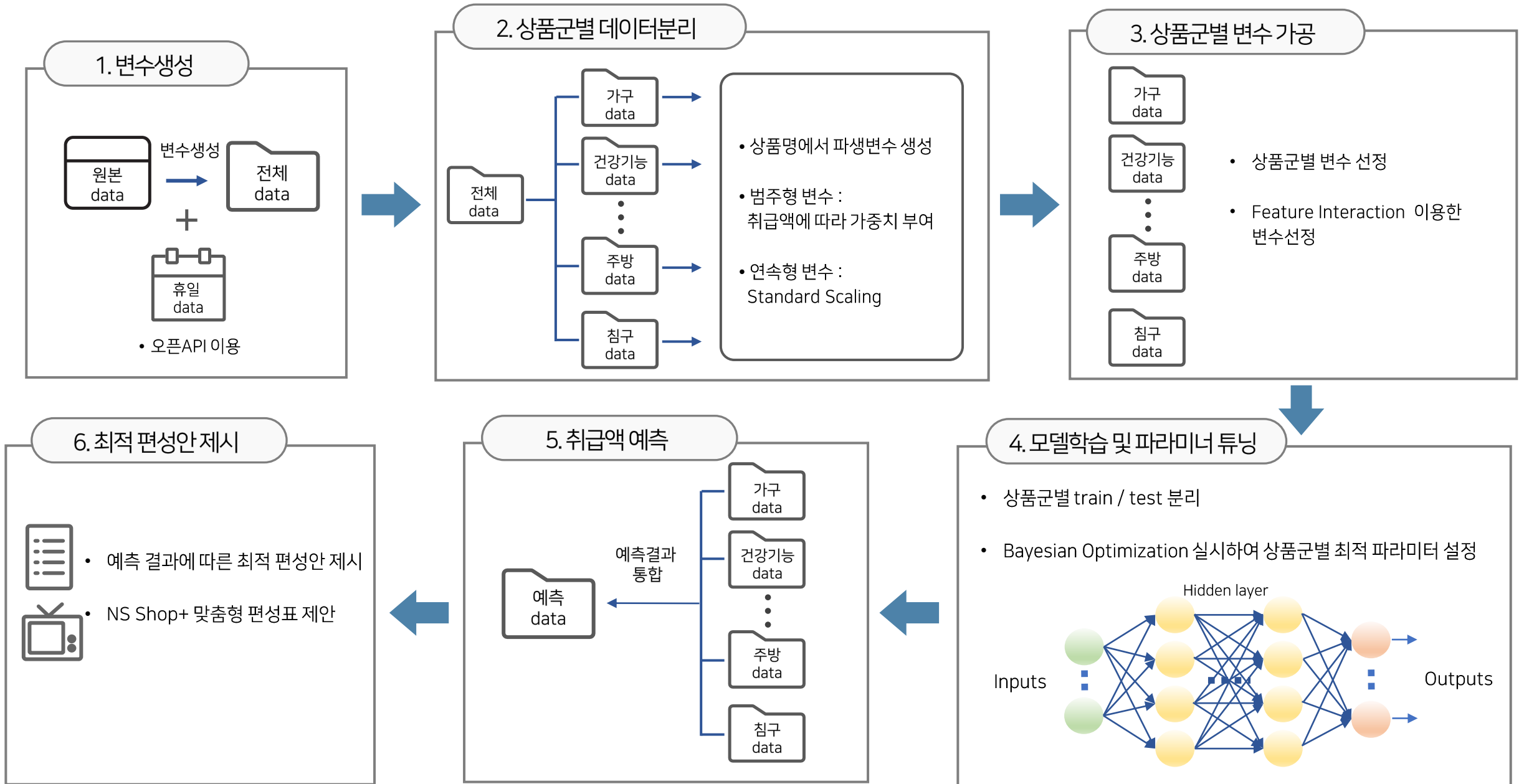
✓ EDA 결과 & 선행연구 내용에 기반

→ **상품군별 개별 모델 구축**



1. 각 상품군별 구매패턴, 특징 반영

2. '상품명'을 이용한 파생변수 생성 가능 (브랜드, 상품 세부 분류 등)



● 변수 가공

	A	B	C	D	E	F	G	H
1	NS Shop+ 2019.01.01~12.31							
2	방송일시	노출(분)	마더코드	상품코드	상품명	상품군	판매단가	취급액
3	2019-01-01 6:00	20	100346	201072	테이트 남 의류		39,900	2,099,000
4	2019-01-01 6:00		100346	201079	테이트 여 의류		39,900	4,371,000
5	2019-01-01 6:20	20	100346	201072	테이트 남 의류		39,900	3,262,000
6	2019-01-01 6:20		100346	201079	테이트 여 의류		39,900	6,955,000
7	2019-01-01 6:40	20	100346	201072	테이트 남 의류		39,900	6,672,000
8	2019-01-01 6:40		100346	201079	테이트 여 의류		39,900	9,337,000
9	2019-01-01 7:00	20	100305	200974	오모떼 레속옷		59,000	6,819,000
10	2019-01-01 7:20	20	100305	200974	오모떼 레속옷		59,000	15,689,000
11	2019-01-01 7:40	20	100305	200974	오모떼 레속옷		59,000	25,370,000
12	2019-01-01 8:00	20	100808	202377	CERINI by 의류		59,900	16,133,000
13	2019-01-01 8:20	20	100808	202377	CERINI by 의류		59,900	30,061,000
14	2019-01-01 8:40	20	100808	202377	CERINI by 의류		59,900	53,542,000

✓ '방송일시' 는 취급액에 영향을 미치는 핵심적인 변수

→ '요일', '주말', '분기', '계절' 변수 추가

→ 오픈 API 이용하여 '공휴일', '연속 휴일' 여부 나타내는 변수 추가

✓ '방송일시' 에서 시간대는 이후 상품군별로 분류 후 파생변수로 가공

✓ '노출(분)' 의 결측치를 처리하고, 연속해서 2회 이상 방송할 경우 '연속방영 가중치' 생성

✓ 구매에 영향을 줄 수 있다고 판단한 '월급일' 변수를 추가(10, 17, 20, 25일)

● 변수 제외

✓ 시청률 데이터

- 취급액과 관계 파악되지 않음 → 모바일 이용률 증가가 원인으로 파악
- 2020년 6월 시청률 데이터가 없으므로, 학습에 사용할 수 없음 → 학습 과정에서 제외

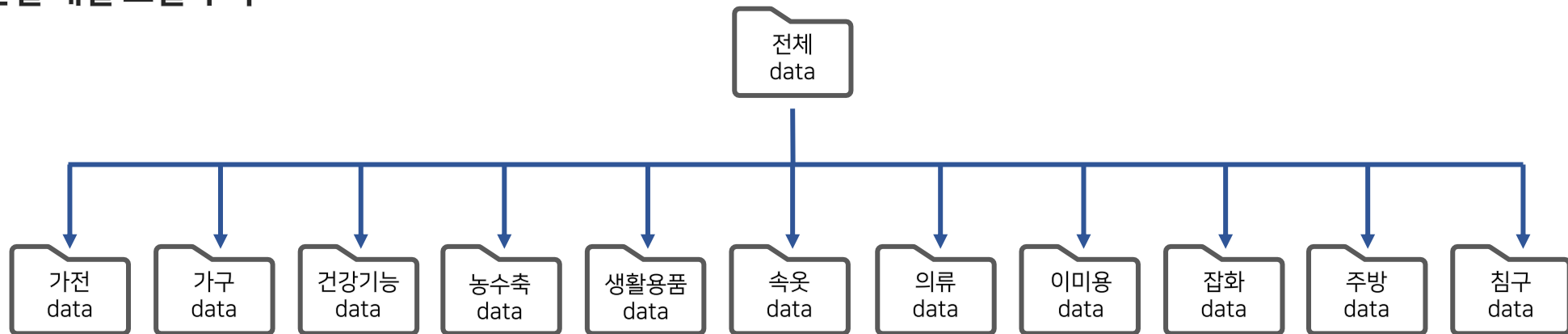
✓ 날씨 데이터

- 가정에서 TV 시청 비율 감소하면서 기상조건의 영향력 감소
- 방송 편성 시 미래 시점의 정확한 기상 데이터 확보 어려움
- 기상예보를 활용할 수 있으나, 정확도 낮음
- 기상 데이터는 전통적으로 홈쇼핑 취급액 예측에서 중요한 변수이지만, 본 분석에서는 제외

✓ 마더코드, 상품코드

- 협력사, 상품명, 카테고리과 관계없이 임의로 생성되는 코드
- 학습 과정에서 제외

● 상품군별 개별 모델 구축



- ✓ 각 상품군을 기준으로 11개 데이터로 분리하여 모델 구축 → 각 상품군의 특징, 패턴을 반영하여 학습
- ✓ 상품군별로 모델을 구축함으로써 '상품명'에서 새로운 파생변수 생성 → 브랜드, 세부 분류, 결제 방식 등

● 가전 상품군 파생변수 생성 예시

상품명
일시불 LG 통돌이 세탁기
무이자 삼성 노트북 9 메탈 기본형
무이자 삼성 유로피안 냉장고
일시불 캐리어 벽걸이에어컨 6형



브랜드	세부분류	결제방식
LG	세탁기	일시불
삼성	노트북	할부
삼성	냉장고	할부
캐리어	에어컨	일시불

● 상품명 파생변수 키워드 (예시)

가전

브랜드		상품종류		결제방식	
키워드	변수	키워드	변수	키워드	변수
LG	LG	세탁기	세탁기	일시불	일시불
삼성	삼성	냉장고	냉장고	무이자	할부
캐리어, 클라세	기타	TV	TV	해당없음	기타
		노트북	노트북		
		공기청정기	공기청정기		
		청소기	청소기		
		건조기	건조기		
		에어컨	에어컨		

농수축

상품종류			
키워드	변수		
곱창, 아구채, 장조림, 배즙, 요일, 두유, 곡물그대로, 철판, 재첩국, 젓갈, 김치, 강정, 머루원액, 갈비탕, 부각, 떡, 쥐포, 식탁앤김, 양갱, 소한마리, 선식, 국밥, 낙곱새, 곤드레, 매실, 스테이크, 닭발, 감말랭이, 콩국수, 호빵	가공		
옥수수, 아카시아꿀, 샤인머스켓, 쌀, 시래기, 사과, 고구마, 고춧가루, 두부, 한라봉, 메주, 오렌지	농산		
수산물, 생선, 옥돔, 굴비, 참굴, 가자미, 갑오징어, 오징어, 꼬막, 새우, 홍어, 골뱅이, 조기, 문어, 꽃게, 장어, 고등어, 전복, 대구, 랍스터, 매생이, 갈치, 낙지, 다시, 과메기, 황태	수산		
소갈비살, LA갈비, 오리, 돼지, 치마살, 한우 한판, 우삼겹, 안창살	축산		

가구

브랜드		상품종류		결제방식	
키워드	변수	키워드	변수	키워드	변수
보루네오	보루네오	침대	침대	(일)	일시불
삼익	삼익	소파, 리클라이너, 휴카우치, 행복돌침대	소파	(무)	할부
장수	장수			해당없음	기타
한샘	한샘				
이누스바스	이누스바스	하이바스, 이누스바스	욕실		
이조농방	이조농방				
뉴벨라홈	뉴벨라홈	불박이장	불박이장		
레스토닉	레스토닉	해당없음	기타		

생활용품

상품종류					
키워드	변수	키워드	변수	키워드	변수
안마의자	안마의자	마사지, 손안마기	마사지	선풍기, 서큘레이터, 에어컨, 원쿨레이터, 에어컨	냉방
의자	의자	매트릭스,메모리폼	매트릭스		
이불브	금고	청소기,로봇청소기	청소기		
마스크	마스크	매트, 카페트매트, 온열매트, 온수매트, 물매트	매트	세제, 섬유유연제, 건조대, 빨래살통, 화이트업	세탁
샤워기	욕실				
비데	비데				
벽지, 단열벽지	벽지	통증패치, 수액패치, 동전패치	패치	플레티늄, 스쿼트머신, 워킹머신	운동기구
옷걸이, 리빙박스	생활용품				
테이블, 교자상, 의자	가구				

건강기능

브랜드		상품종류		결제방식	
키워드	변수	키워드	변수	키워드	변수
종근당	종근당	루테인	루테인	(일), 일시불	일시불
광동	광동	석류	석류	(무), 무이자	할부
뉴트리원	뉴트리원	구미	구미	해당없음	기타
해당없음	기타	홍삼	홍삼		
		프리바이오틱스, 유산균	유산균		

주방

브랜드				상품종류	
키워드	변수	키워드	변수	키워드	변수
락앤락, 제오닉	락앤제오	한일	한일	칼, 채칼	칼
램프쿡	램프쿡	쿠쿠	쿠쿠	밀폐용기	밀폐용기
린나이	린나이	PN	PN	냄비	냄비
매직쉐프	매직쉐프	프로피쿡	프로피쿡	압력밥솥, 압력솥	밥솥
올리고	올리고	다지기, 블렌더, 믹서기	믹서기	가스레인지	가스레인지
am마카롱	am마카롱	프라이팬, 후라이팬	후라이팬	전자레인지	전자레인지
쿠진나이프	쿠진나이프	살균건조기	살균건조기	그릴레인지	그릴레인지
파뮈에, 독일지니어스	파뮈독일	에어프라이어	에어프라이어	텀블러	텀블러

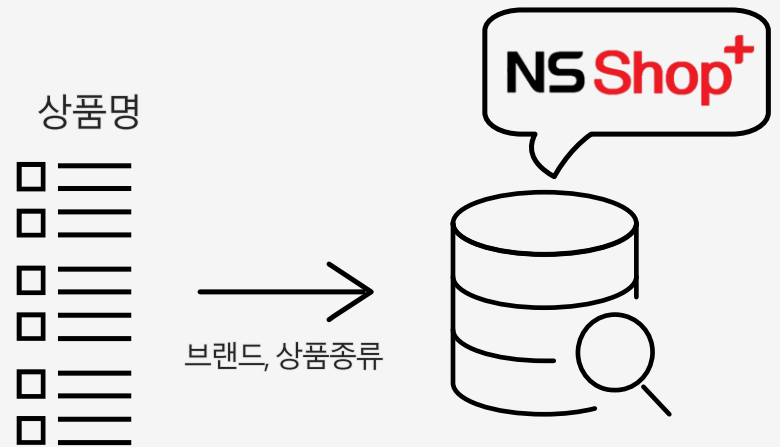
- 상품명 자동 분류 시 문제점

- 형태소분석기를 이용하여 상품명을 자동 분류할 경우의 문제

- ① NS 홈쇼핑은 다양한 중소기업들의 상품을 취급하므로, 브랜드명, 상품종류 등의 변수 차원이 지나치게 증가
- ② 일부 대기업을 제외하고 동일한 브랜드나 상품종류에 대한 충분한 학습 데이터 수집 어려움
- ③ 상품 정보를 정확하기 추출하기 어려운 상품명 존재
(ex "종근당건강 전립선 프라임", "비버리힐스폴로클럽 남성기초세트", "참존탐뉴스 지이링클 세트")

✓ 상품군별 상품명 활용을 위한

→ NS Shop+ 상품명 "키워드사전" 구축



- 단순히 기계적인 형태소 분류를 통해 학습을 실시하는 것이 아니라, 정성적인 분류를 통하여 키워드 사전 구축
- 기존 데이터를 사용하여 키워드 사전 생성
→ 새로운 상품 편성 시 키워드 사전에 단어를 추가하여 학습 시 사용

● 추가 변수 전처리

- 브랜드, 상품군 등 범주형 변수는 one-hot encoding 실시할 경우

데이터 차원이 지나치게 증가하는 문제 발생

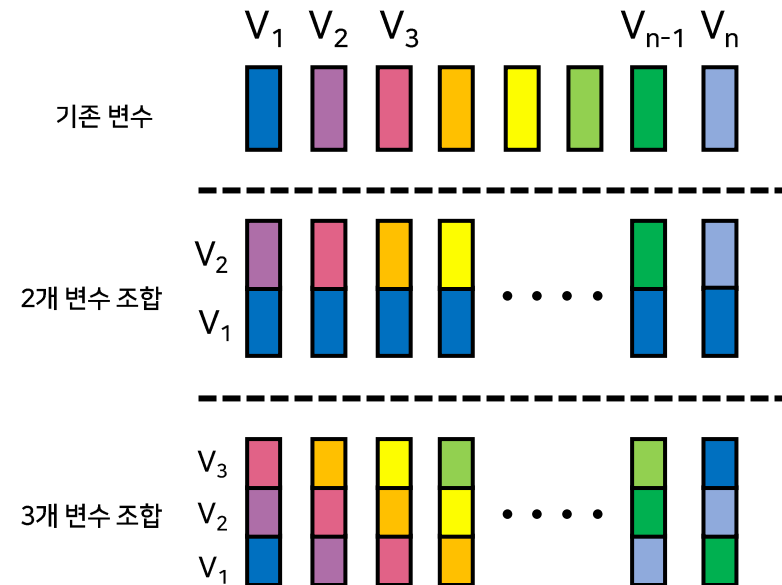
→ 취급액을 기준으로 범주형 변수들을 scaling 하여 수치형 변수로 변환

- 기타 수치형 변수들에 대해서도 모두 Standard Scaling 실시

● 변수 선정

✓ Feature interaction : 변수 추가생성

- ① 취급액과의 상관관계가 0 이하인 변수 → 해당 변수 값을 역수 + 로그변환
- ② 상관관계가 0.1 이하인 변수 삭제
- ③ 변수 2개, 3개를 더하여 새로운 변수조합 생성
- ④ 기존 개별 변수보다 취급액과의 상관관계가 낮아지거나 0.4 이하인 변수 삭제



● 사용 알고리즘 - 심층신경망(DNN)

..... [심층신경망 선정 배경]

- DNN과 GBDT(Gradient boosting)이 LR과 SVR에 비해 성과가 좋았으며,

DNN의 경우 GBDT보다 적은 데이터로 학습이 가능함

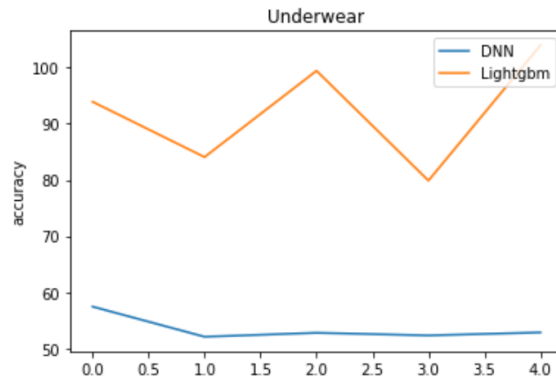
(Hung, Chen-Ying, et al, 2017)

- ANN과 Tree 기반의 모델을 비교했을 때, 회귀의 예측성과에서 ANN이 높은 성과를 보임

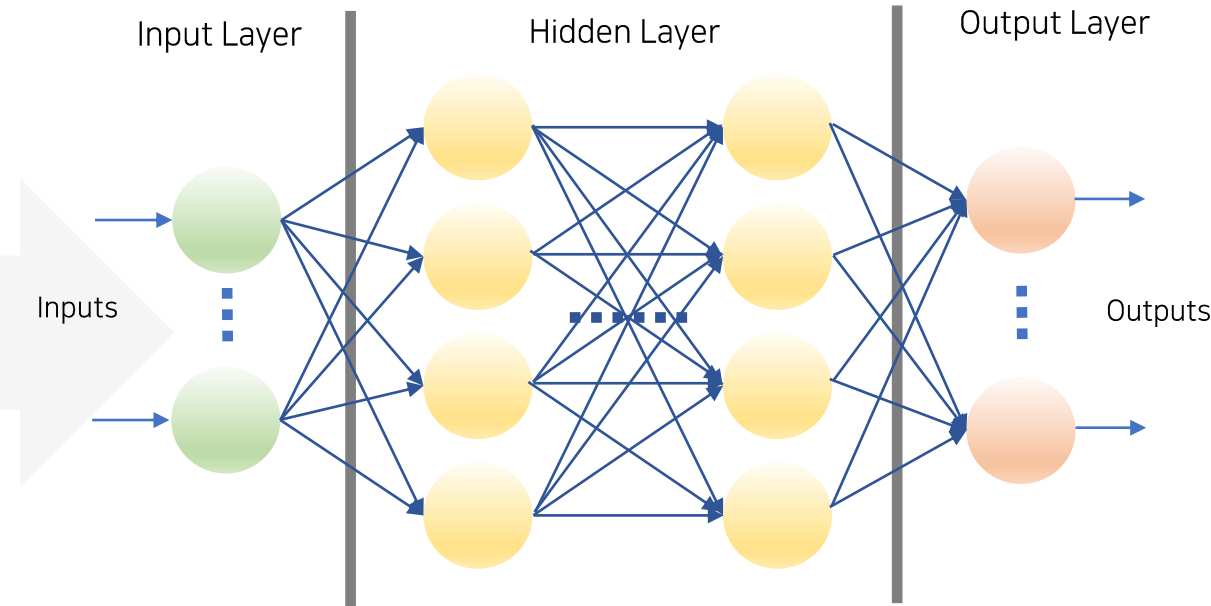
(Ahmad, Muhammad Waseem, Monjur Mourshed, and Yacine Rezgui, 2017)

- 또한 본 데이터를 이용하여 구축한 상품군별 모델에서 여러 차례 반복 실험한 결과,

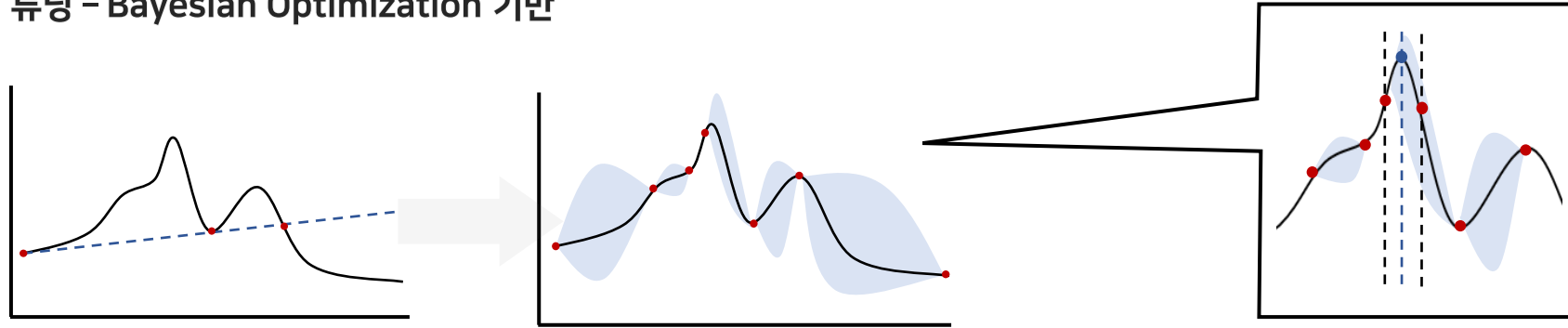
선형 회귀 / 트리 기반 회귀 알고리즘 보다, 심층 신경망의 예측 오차가 적고 안정성이 높았음



● 심층신경망 구조도



● DNN 파라미터 튜닝 - Bayesian Optimization 기반



✓ Bayesian Optimization → 최적 파라미터 선정

- 베이지안 최적화는 지역최적화를 방지에 효과적인 전략이다(Mockus et al., 1978)
- 베이지안 최적화를 사용하면 전문가 지식 수준의 성과를 달성할 수 있다(J. Snoek et al., 2012)

✓ 베이지안 최적화 값을 기준으로 Manual search 실시
→ 국소 최적 파라미터 탐색

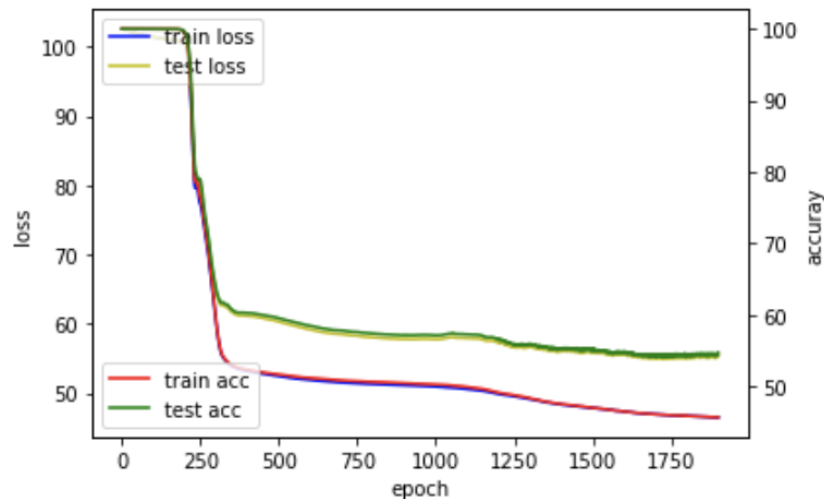
● 상품군별 최적 파라미터 탐색 결과

	농수축 (agriculture)	가구 (furniture)	속옷 (underwear)	이미용 (beauty)	생활용품 (life)	의류 (clothing)	잡화 (accessory)	침구 (bedding)	가전 (appliance)	건강기능 (health)	주방 (kitchen)
Layer1	16	28	56	32	36	40	16	24	48	60	44
Layer2	24	42	84	48	54	60	24	36	72	90	66
Layer3	24	42	84	48	54	60	24	36	72	90	66
Layer4	32	56	112	64	72	80	32	48	96	120	88
Layer5	36	63	126	72	81	90	36	54	108	135	99
Layer6	36	63	126	72	81	90	36	54	108	135	99
Layer7	32	56	112	64	72	80	32	48	96	120	88
Layer8	24	42	84	48	54	60	24	36	72	90	66
Layer9	24	42	84	48	54	60	24	36	72	90	66
Layer10	12	21	42	24	27	30	12	18	36	90	33

● 모델별 MAPE

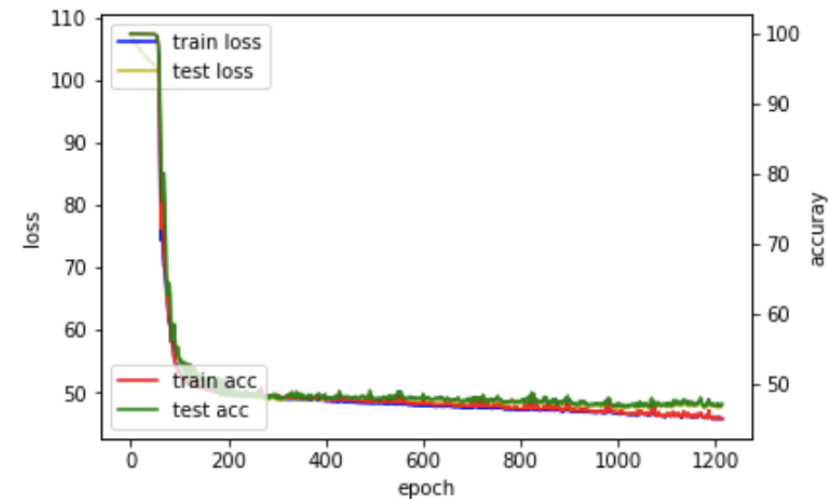
상품군 모델별로 학습 후 성과 확인

잡화(accessory)



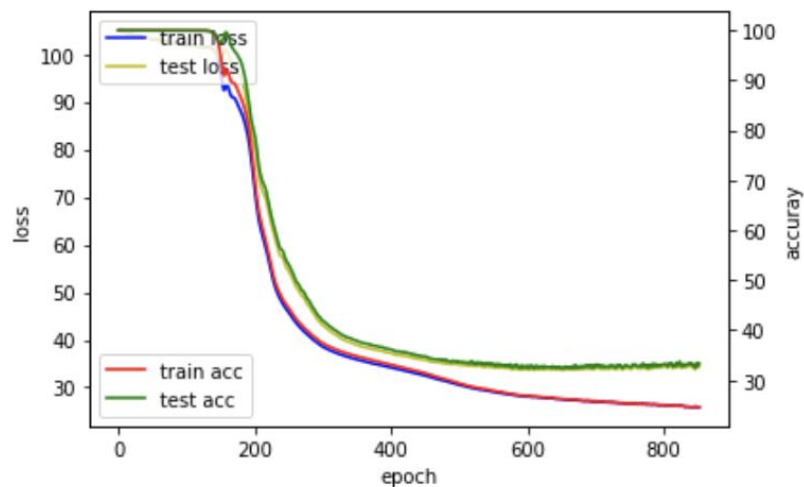
Train : 48.99 / Test : 49.86

가전(appliance)



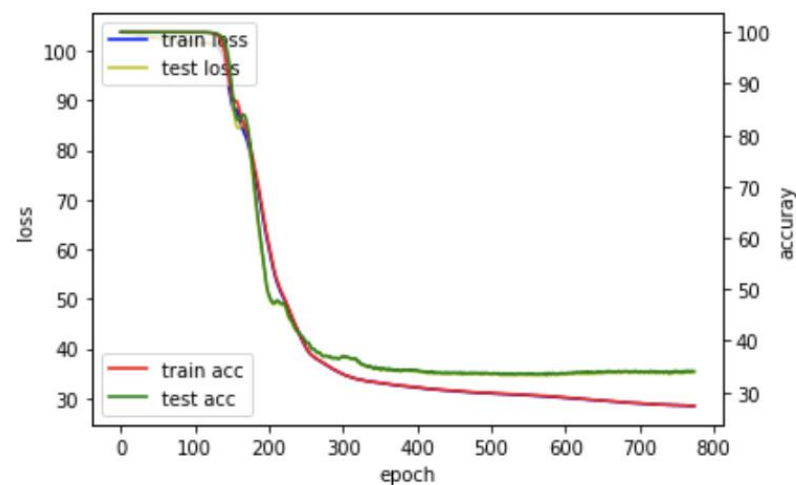
Train : 45.12 / Test : 46.48

이미용(beauty)



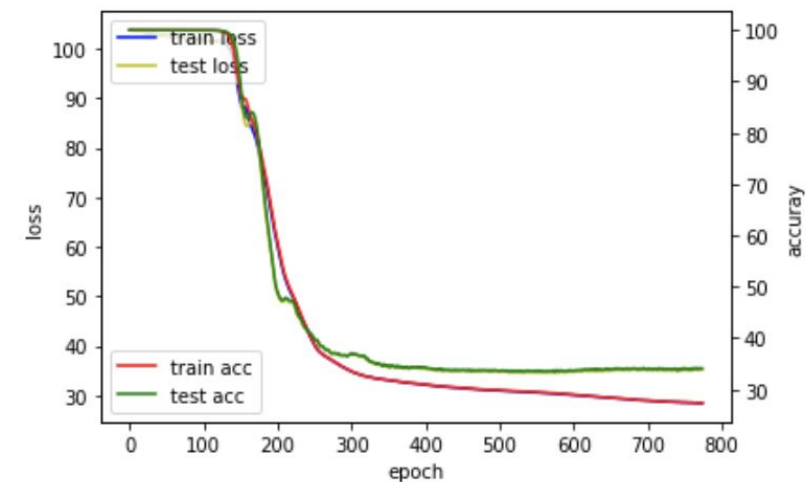
Train : 22.61 / Test : 31.35

침구(bedding)



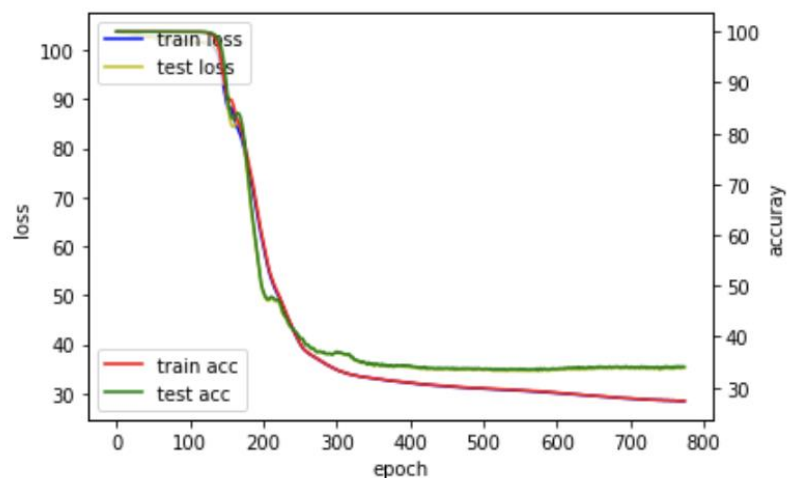
Train : 24.55 / Test : 32.87

의류(clothing)



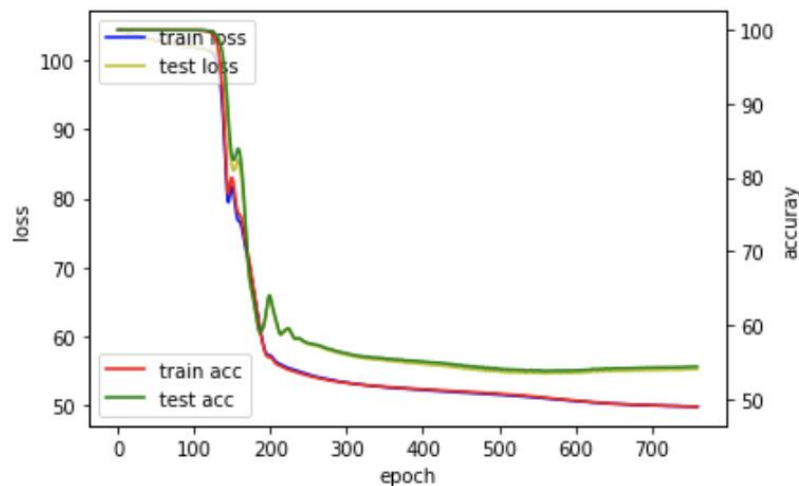
Train : 34.46 / Test : 41.81

농수축(agriculture)



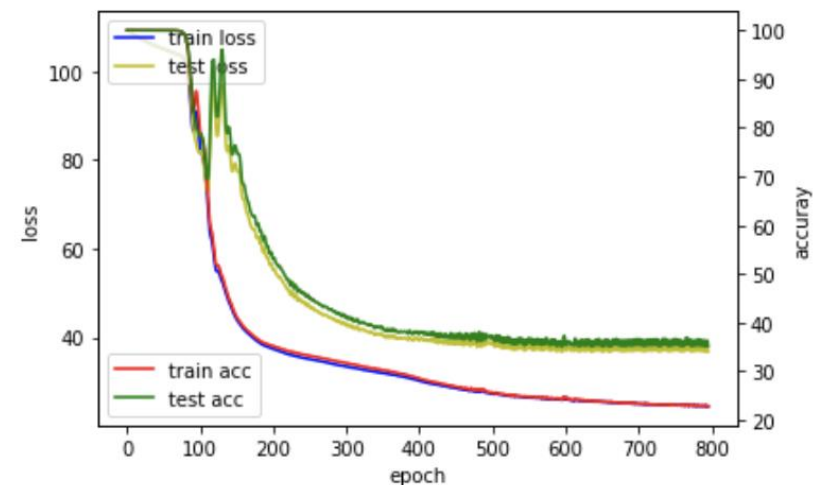
Train : 27.53 / Test : 30.10

가구(furniture)



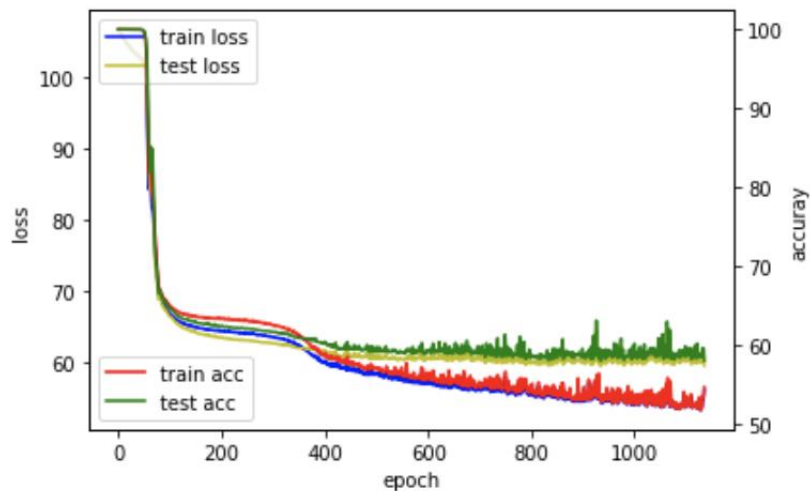
Train : 49.18 / Test : 53.96

건강기능(health)



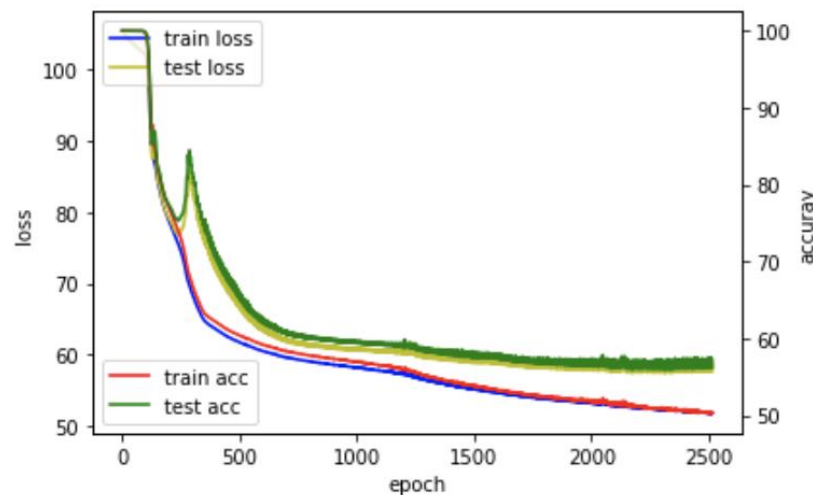
Train : 21.54 / Test : 33.83

주방(kitchen)



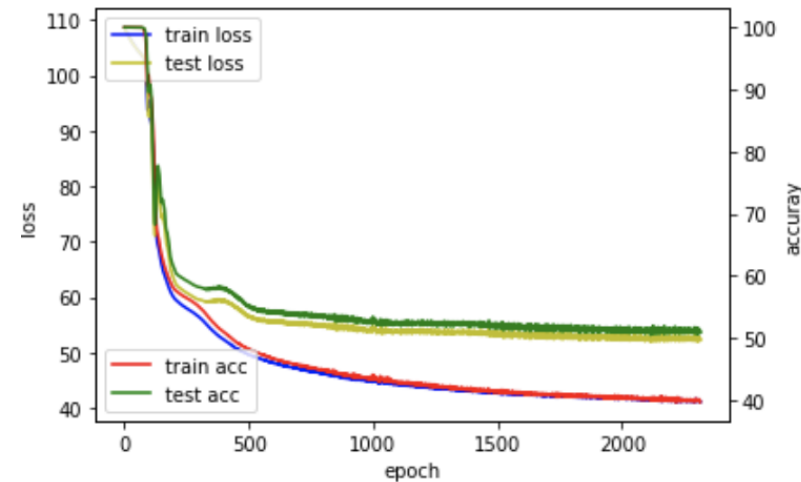
Train : 46.92 / Test : 48.66

생활용품(life)



Train : 51.32 / Test : 53.85

속옷(underwear)



Train : 40.80 / Test : 49.59

● 최적 편성표 제안 구도

NS Shop+

- 기존 20년 6월 편성표를 이용하여 상품군별 모델을 사용하여 예측 실시
- 예측값과 작년 취급액평균 비교

취급액 예측값



동일 시간대 · 상품군의
2019년 평균 취급액

① 예측값 > 2019년 평균 취급액

- 올바르게 상품이 편성이 된 것으로 판단, 기존 편성 항목 유지

② 예측값 < 2019년 평균 취급액

- 편성표 수정을 통한 취급액 증대 가능
- **작년 동시간대 평균 취급액이 높은 상품군의 상품으로 재편성**
- 재편성할 구체적인 상품 제시도 가능하지만, 해당시점의 NS홈쇼핑의 협력업체 상황, 계약관계, 내부상황 등에 따라 판매가능 상품이 달라짐
- **개별 상품이 아닌 상품군으로 재편성 품목 제시**

③ 재편성 품목 검증

- 재편성된 상품군의 데이터를 다시 모델에 입력하여 취급액 증가 여부 검증
- 취급액이 증가하지 않을 경우, 다른 상품군으로 대체과정 반복



● 최적 편성표 구축 예시

[20년 6월 편성표]

방송일시	상품군	상품명
2020-06-01 06:00	가전	LG전자
2020-06-01 06:20	의류	벨레즈온
⋮	⋮	⋮
2020-06-31 23:40	가구	보루네오

예측 실시

[예측 취급액]

단위 : 만원

예측값
5930
1973
⋮
1772

[19년 동일 상품군의 평균 취급액]

단위 : 만원

상품군별 평균 취급액
2495
3148
⋮
1500

비교

[19년 시간대별 상품군별 평균 취급액]

단위 : 만원

방송일시 \ 상품군	농수축	가구	속옷	이미용	생활용품	의류	잡화	침구	가전	건강기능	주방
2020-06-01 06:20	3221	2123	2214	5823	1009	4148	1321	1412	3201	1105	808

예측값이 전년도 동일 상품군 평균취급액 보다 낮으므로,
전년도 해당 시간대와 상품군에서 평균 취급액이 가장 높은
'이미용' 상품군의 상품으로 재편성

- ✓ 이미용 상품군에서 판매가능한 상품을 모델에 입력하여 취급액 증가 확인
- ✓ 취급액 증가가 예측되지 않을 경우, 동일 상품군 내 다른 상품으로 편성하거나,
두 번째로 평균 취급액이 높은 '의류' 상품군으로 대체하여 편성표 수정

2020 빅 콘테스트 데이터분석 분야 챔피언리그

감사합니다



● NS Shop+ 상품별 키워드 사전 목록

가전

브랜드		상품종류		결제방식	
키워드	변수	키워드	변수	키워드	변수
LG	LG	세탁기	세탁기	일시불	일시불
삼성	삼성	냉장고	냉장고	무이자	할부
캐리어, 클라세	기타	TV	TV	해당없음	기타
		노트북	노트북		
		공기청정기	공기청정기		
		청소기	청소기		
		건조기	건조기		
		에어컨	에어컨		

농수축

상품종류			
키워드	변수		
곱창, 아구채, 장조림, 배즙, 요일, 두유, 곡물그대로, 철판, 재첩국, 젓갈, 김치, 강정, 머루원액, 갈비탕, 부각, 떡, 쥐포, 식탁앤김, 양갱, 소한마리, 선식, 국밥, 낙곱새, 곤드레, 매실, 스테이크, 닭발, 감말랭이, 콩국수, 호빵	가공		
옥수수, 아카시아꿀, 샤인머스켓, 쌀, 시래기, 사과, 고구마, 고춧가루, 두부, 한라봉, 메주, 오렌지	농산		
수산물, 생선, 옥돔, 굴비, 참굴, 가자미, 갑오징어, 오징어, 꼬막, 새우, 홍어, 굴뱅이, 조기, 문어, 꽃게, 장어, 고등어, 전복, 대구, 랍스터, 매생이, 갈치, 낙지, 다시, 과메기, 황태	수산		
소갈비살, LA갈비, 오리, 돼지, 치마살, 한우 한판, 우삼겹, 안창살	축산		

가구

브랜드		상품종류		결제방식	
키워드	변수	키워드	변수	키워드	변수
보루네오	보루네오	침대	침대	(일)	일시불
삼익	삼익	소파, 리클라이너, 휴카우치, 행복돌침대	소파	(무)	할부
장수	장수			해당없음	기타
한샘	한샘				
이누스바스	이누스바스	하이바스, 이누스바스	욕실		
이조농방	이조농방				
뉴벨라홈	뉴벨라홈	불박이장	불박이장		
레스토닉	레스토닉	해당없음	기타		

생활용품

상품종류					
키워드	변수	키워드	변수	키워드	변수
안마의자	안마의자	마사지, 손안마기	마사지	선풍기, 서큘레이터, 써큘레이터, 에어컨, 원쿨레이터, 에어컨러	냉방
의자	의자	매트릭스,메모리폼	매트릭스		
이불브	금고	청소기, 로봇청소기	청소기		
마스크	마스크	매트, 카페트매트, 온열매트, 온수매트, 물매트	매트	세제, 섬유유연제, 건조대, 빨래살통, 화이트업	세탁
샤워기	욕실				
비데	비데				
벽지, 단열벽지	벽지	통증패치, 수액패치, 동전패치	패치	플레티늄, 스쿼트머신, 워킹머신	운동기구
옷걸이, 리빙박스	생활용품				
테이블, 교자상, 의자	가구				

건강기능

브랜드		상품종류		결제방식	
키워드	변수	키워드	변수	키워드	변수
종근당	종근당	루테인	루테인	(일), 일시불	일시불
광동	광동	석류	석류	(무), 무이자	할부
뉴트리원	뉴트리원	구미	구미	해당없음	기타
해당없음	기타	홍삼	홍삼		
		프리바이오틱스, 유산균	유산균		

주방

브랜드				상품종류	
키워드	변수	키워드	변수	키워드	변수
락앤락, 제오닉	락앤제오	한일	한일	칼, 채칼	칼
램프쿡	램프쿡	쿠쿠	쿠쿠	밀폐용기	밀폐용기
린나이	린나이	PN	PN	냄비	냄비
매직쉐프	매직쉐프	프로피쿡	프로피쿡	압력밥솥, 압력솥	밥솥
올리고	올리고	다지기, 블렌더, 믹서기	믹서기	가스레인지	가스레인지
am마카롱	am마카롱	프라이팬, 후라이팬	후라이팬	전자레인지	전자레인지
쿠진나이프	쿠진나이프	살균건조기	살균건조기	그릴레인지	그릴레인지
파뮈에, 독일지니어스	파뮈독일	에어프라이어	에어프라이어	텀블러	텀블러

의류

브랜드				상품종류	
키워드	변수	키워드	변수	키워드	변수
디즈니	디즈니	아문센	아문센	밍크	밍크
유리진	유리진	아주아	아주아	수트	정장
페플럼제이	페플럼제이	알렉스하운드	알렉스하운드	트랙수트, 멀티웨어	운동
헤스티지	헤스티지	어반시크릿	어반시크릿	코트, 롱코트, 벤치코트	코트
EXR	EXR	에르나벨	에르나벨	패딩, 구스다운	패딩
K-SWISS	K-SWISS	엔셀라두스	엔셀라두스	자켓, 재킷, 윈드브레이크, 핫멜트, 니트점퍼	자켓
USPA	USPA	오렐리안	오렐리안	베스트	베스트
마리노블	마리노블	이동수골프	이동수골프	블라우스	블라우스
그렉노먼	그렉노먼	임페리얼	임페리얼	셔츠	셔츠
더블유베일	더블유베일	코몽트	코몽트	터틀넥, 니트, 스웨터	니트
도네이	도네이	코펜하겐럭스	코펜하겐럭스	데님팬츠, 데님	청바지
디베이지	디베이지	크리스티나앤코	크리스티나앤코	밴딩팬츠, 팬츠, 라인핏, 레깅스	바지
디키즈	디키즈	타운젠트	타운젠트	무스탕	무스탕
라라쎄	라라쎄	테이트	테이트	폴코디, 앙상블	세트
레드캠프	레드캠프	팜스프링스	팜스프링스	원피스	원피스
헤비추얼	헤비추얼	PAT	PAT	맨투맨, 티셔츠, 약기모티, 모크넥	티셔츠
헤스티지	헤스티지	NNF	NNF		
대동모피	대동모피	젠트웰	젠트웰		

속옷

브랜드		상품종류		성별	
키워드	변수	키워드	변수	키워드	변수
헤드, 프로스펙스, 카파, 푸마	스포츠	패키지, 반팔상하세트, 세트	세트	여성, 여아, 이너웨어, 슬리머, 매직니퍼팬티, 거들팬티, 보정런닝, 이지탑, 란주, 웨이퍼, 슬리머, 브라, 브라탑, 브라렛, 보정브라, 브라팬티, 이지웨어, 원피스, 이너웨어, 레깅스, 카파 런닝 패키지, 푸마 드라이셀라인 뉴 런닝 9종	여성
에버라스트, 보디가드	에버보디	언더셔츠	언더셔츠		
실크트리, 자연감성	실크자연	드로즈, 트렁크	남성팬티		
라쉬반	라쉬반	매직니퍼팬티, 거들팬티	여성팬티		
레이프릴	레이프릴	보정런닝, 이지탑		남성, 머슬, 남아, 언더셔츠, 드로즈, 트렁크, 런닝, 언더탑	남성
루시헨느	루시헨느	런닝, 언더탑			
몬테밀라노	몬테밀라노	란주, 웨이퍼, 슬리머			
발레리	발레리	브라, 브라탑, 브라렛, 보정브라			

침구류

브랜드		상품종류	
키워드	변수	키워드	변수
보몽드	보몽드	침구 풀세트, 침구세트	침구
한샘, 지나송, 한스데코	한샘지나		
리앤코리아	리앤코리아	커튼	커튼
한빛	한빛	카페트	카페트
효재	효재		

잡화

브랜드		상품 유형	
키워드	변수	키워드	변수
루이띠에	루이띠에	목걸이	목걸이
구찌	구찌	팔찌	팔찌
기라로쉬	기라로쉬	반지	반지
레노마, 레비노블	레노레비	선글라스	선글라스
버버리	버버리	지갑	지갑
로베르타, 엘르, 삭루츠, 칼리베이직, 도스문도스, 소노비	로베엘르	투웨이백, 크로스백, 토트백, 버킷백, 숄더백, 클러치백, 백팩, 사첼, 체인 숄더, 패브릭백, 카메라백, 체인백,	가방
시스마르스	시스마르스		
아가타	아가타		
에버라스트	에버라스트	스카프	스카프
엘르	엘르	양산	양산
오델로	오델로	스니커즈, 슬립온	스니커즈
월드컵	월드컵	러닝화, 트레킹화, 워킹화	운동화
칼리베이직	칼리베이직	펌프스	펌프스
엘리자베스아덴, 로베르타, 기라로쉬, 페리엘리스	엘리로베	샌들	샌들
프라다, 생로랑, DIOR, 톰포드, 페라가모	프라생로	모자	모자
AAA	AAA	그 외 키워드들	기타
트레스패스, 레코바, W클라우드	트레레코		

이미용

브랜드		상품 유형	
키워드	변수	키워드	변수
보닌	보닌	기초세트, 에센스, 선스틱, 썬스틱, 세럼, 크림, 클렌저, 클렌징패드	기초화장품
더블모, 셀럽by재클린, 제니하우스, 프리지아	더셀제프		
아이앤아이, 래쉬톡	아이앤래쉬	파운데이션, 아이라이너, 아이스타일러, 쿠션, 커버팩트	화장품
시크릿, 바비리스	시크바비	속눈썹	속눈썹
블링썸	블링썸	샴푸	샴푸
비버리	비버리	틴트, 립스틱	틴트
에이유플러스	에이유플러스	염색, 헤어컬러, 뿌리펌, 셀프뿌리퍼머	염색
엘렌실라, 글레드벨	엘렌글레		
아미니, 족선생	아미족	필링, 에이지큐어, 마스크팩, 마스크, 모공기기관리세트	피부
참존	참존		
바바코코	바바코코	풋샴푸, 풋케어	발관리
프리미엄 클린샤워	노브랜드	샤워, 오일워시	샤워
		젤네일스트립	손톱
		고데기, 고데롤, 볼륨스타일러	헤어기기

● 참고 문헌

- Ahmad, M. W., Mourshed, M., & Rezgui, Y. (2017). Trees vs Neurons: Comparison between random forest and ANN for high-resolution prediction of building energy consumption. *Energy and Buildings*, 147, 77-89.
- Hung, C. Y., Chen, W. C., Lai, P. T., Lin, C. H., & Lee, C. C. (2017, July). Comparing deep neural network and other machine learning algorithms for stroke prediction in a large-scale population-based electronic medical claims database. In *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (pp. 3110-3113). IEEE.
- 김원겸. (2004). 상품유형별 쇼핑과업상황에 따른 소비자 반응에 관한 연구-쇼핑가치의 조절역할을 중심으로. *상품학연구*, 85-107.
- 구명진, & 여정성. (2015). 소비자 정보탐색단계별 탐색어려움과 구매결정-상품 유형성을 중심으로. *소비자정책교육연구*, 11(2), 59-85.
- 정인희. (2015). 소비자 성별에 따른 상품 유형별 관심도 차이, 내재적 혁신성과의 상관관계 및 상품 지각 구조 분석. *한국의류학회지*, 39(4), 505-516.
- Mockus, J., Tiesis, V., & Zilinskas, A. (1978). The application of Bayesian methods for seeking the extremum. *Towards global optimization*, 2(117-129), 2.
- Snoek, J., Larochelle, H., & Adams, R. P. (2012). Practical bayesian optimization of machine learning algorithms. In *Advances in neural information processing systems* (pp. 2951-2959).