

경상북도 데이터 경진대회

추천 알고리즘을 이용한 맞춤형 식품 추천 서비스

박재현 이재성 정은진 최은선

온라인 유통 혁명

온라인 유통은 시간, 공간적 한계를 초월하여
저렴하고 편리한 가치를 소비자에게 전달하면서
최근 5년간 70%나 성장

특히, 코로나19 이후로 비대면 거래 증가로 성장 가속화 예상

온라인쇼핑몰 거래액 규모 추이 (단위: 원)

자료: 통계청 ※온라인몰, 온·오프라인 병행몰 합계



문제 현황

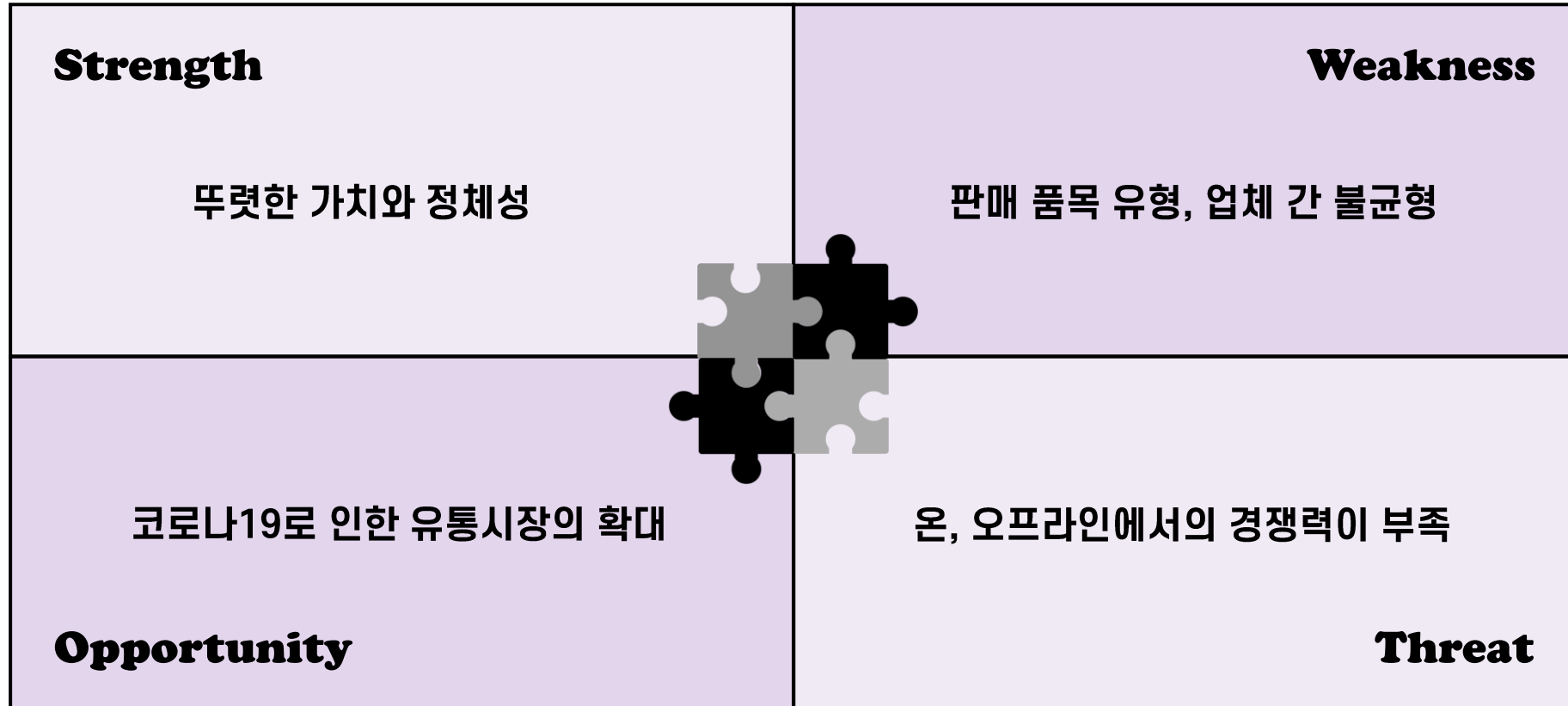
‘농산물 유통약자’를 위한 대책 마련이 시급!

‘농산물 유통개혁’이 온-오프라인을 망라해 전체 농산물 거래 시장에서 진행되어야만 한다.
농민들이 생산한 농산물이 제 값을 받고 팔려 나가는 선순환 시스템을 하루 빨리 마련해야 한다.

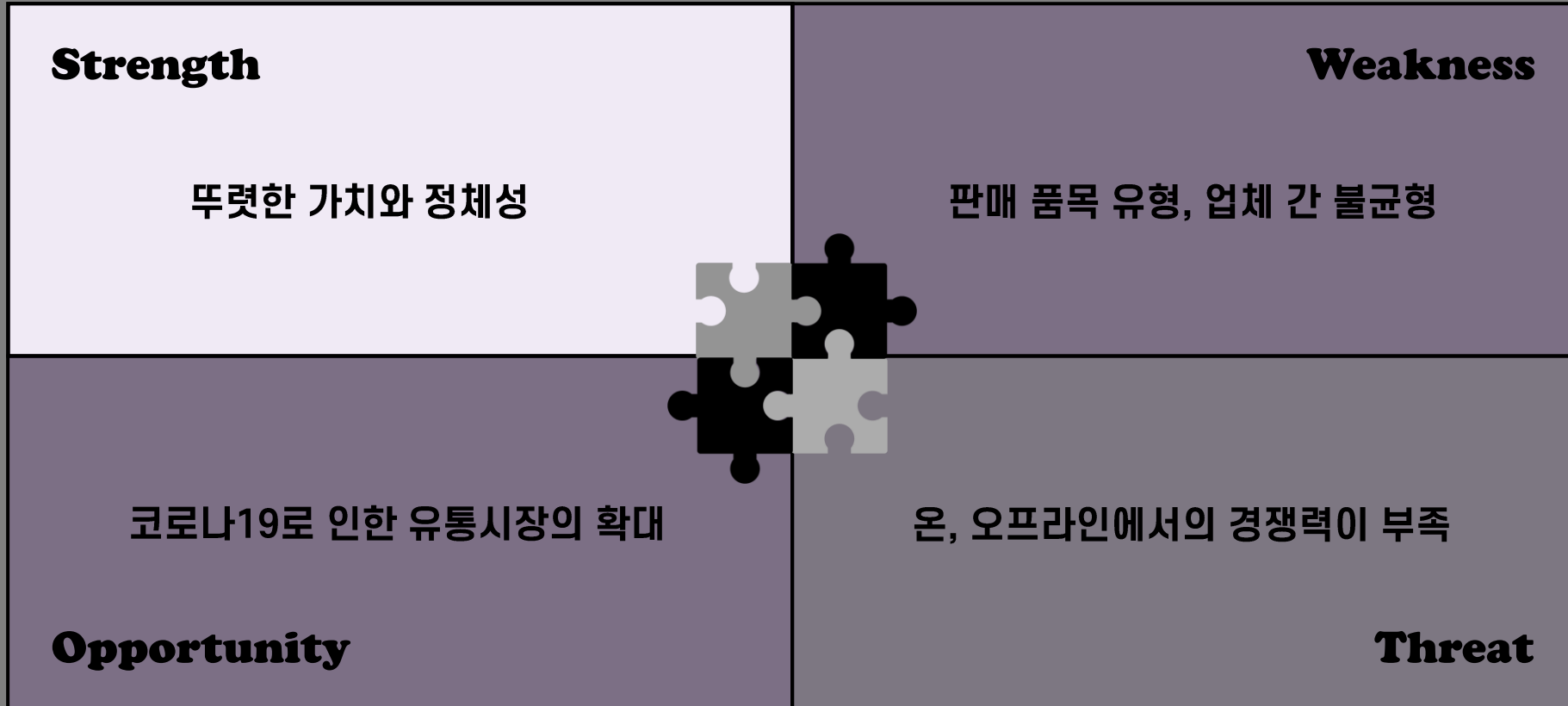
[출처: 이병로, ‘농산물 유통약자’ 위한 대책 마련 시급하다, 한국영농신문]



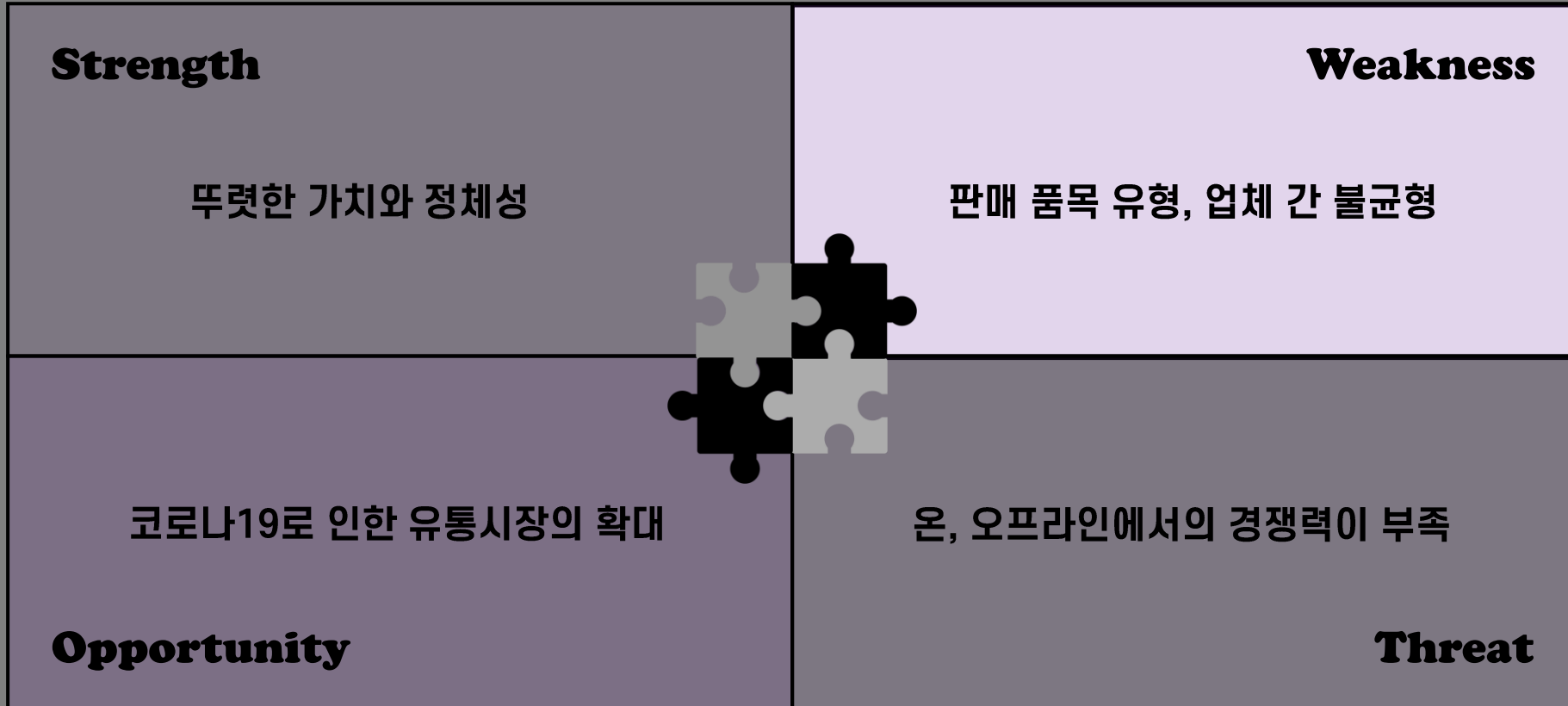
SWOT 분석



SWOT 분석



SWOT 분석



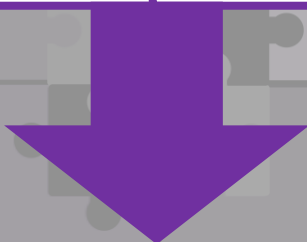
SWOT 분석

판매수량 상위 5개

	업체ID	판매수량
503	baffd9c61cf98685fa3508b7d77bcd087e7dfae0ce34ac...	84
584	dd10b7caf931dcd7bf81505d56531fedd94af61ee698ec...	66
246	5ecdbdeca766d2bc85cd6c76c52221eb5f5c6082647ed8...	65
604	e970449a7fbb26c2239dc5152bf78ac93725a3ad5494af...	62
249	5f90d34c68fb01f90df340c684d12fdb25346c06af0853...	60

판매수량 하위 5개

	업체ID	판매수량
505	bbba8dc3d394f78b9e04f3fc908bef07683d6ac54189d8...	1
227	56d7ad2b9fe1133fd91507e6510a99f81aa90dfaa0ec6f...	1
508	bc741afc77c27ca7748bb1d01e99eddb97aadce7818090...	1
509	bc476e02578c8ee63c4b0c3d125dc9f70a4759efcb3c0...	1
656	ff85e6c9fd60e14b4c2a82a307130075eeca9eb7a2da1d...	1

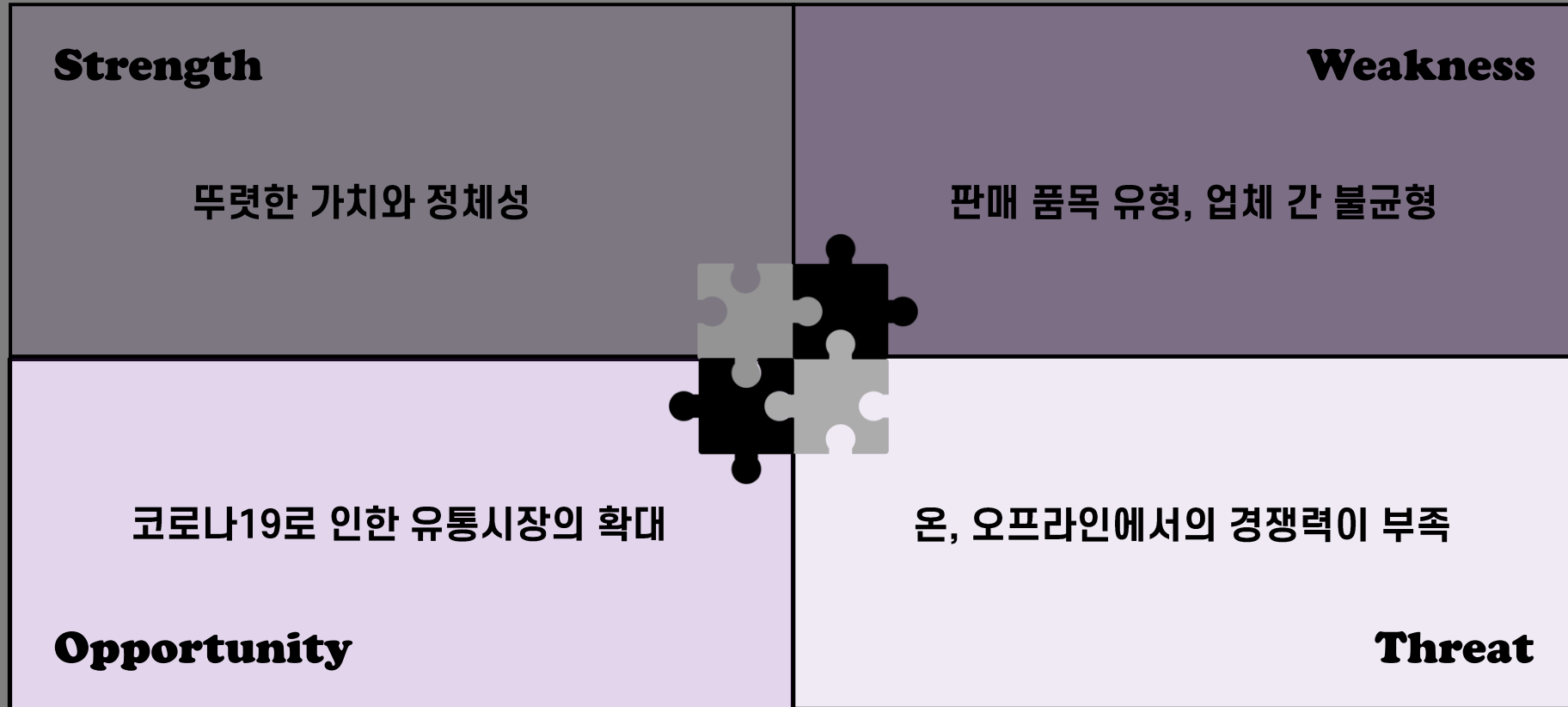


코로나19로 인한 유통시장의 확대

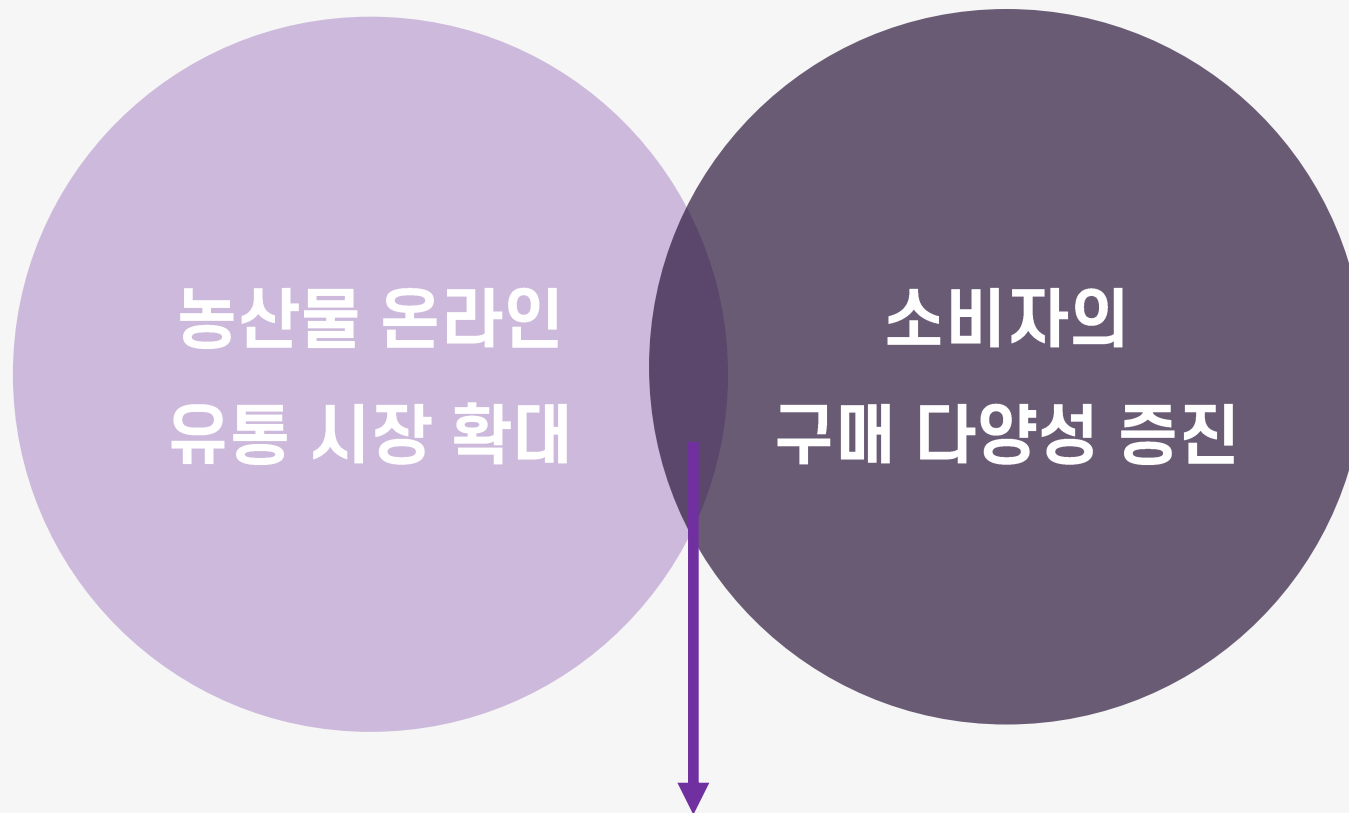
온, 오프라인에서의 경쟁력이 부족

판매수량 상,하위 업체 간 격차가 확연히 나타남

SWOT 분석



문 제 제 기



추천 알고리즘 시스템은 구매전환율의 증가와 다양한 상품의 구매를 하게 하는 효과가 있음

출처: Different but Equal? A Field Experiment on the Impact of Recommendation Systems on Mobile and Personal Computer Channels in Retail (2020)

농산물 온라인

소비자의

입점업체에게는 동등한 기회를, 소비자에게는 다양한 먹거리를 제공하는

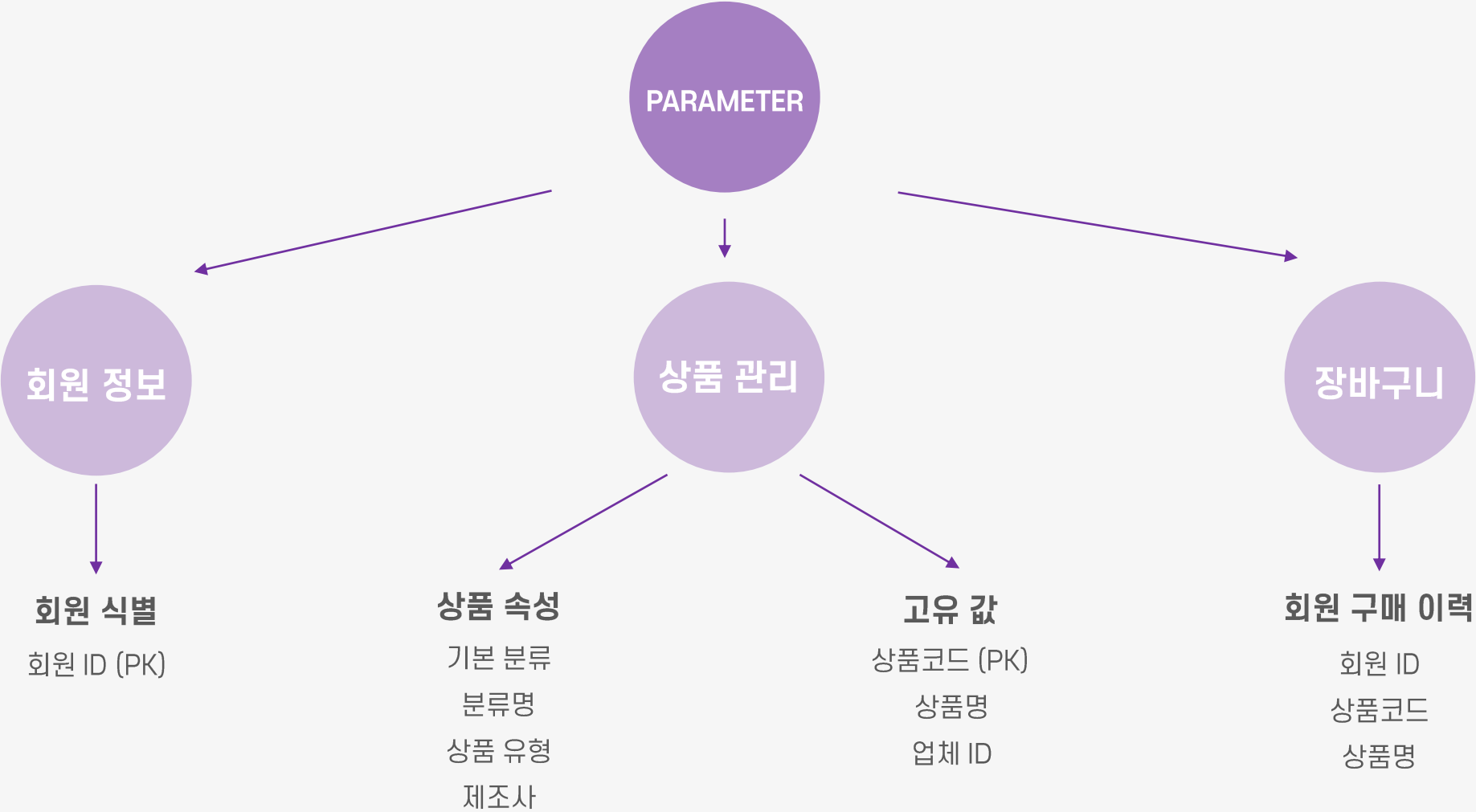
개인 맞춤형 식품 추천 알고리즘



추천 알고리즘 시스템은 구매전환율의 증가와 다양한 상품의 구매를 하게 하는 효과가 있음

출처: Different but Equal? A Field Experiment on the Impact of Recommendation Systems on Mobile and Personal Computer Channels in Retail (2020)

PARAMETER 설정



상품 관리 테이블 전처리

B	C	D
기본분류	2차분류	3차분류
zz	7020	
zz	201030	
zz	7020	
zz	3020	
zz	3020	

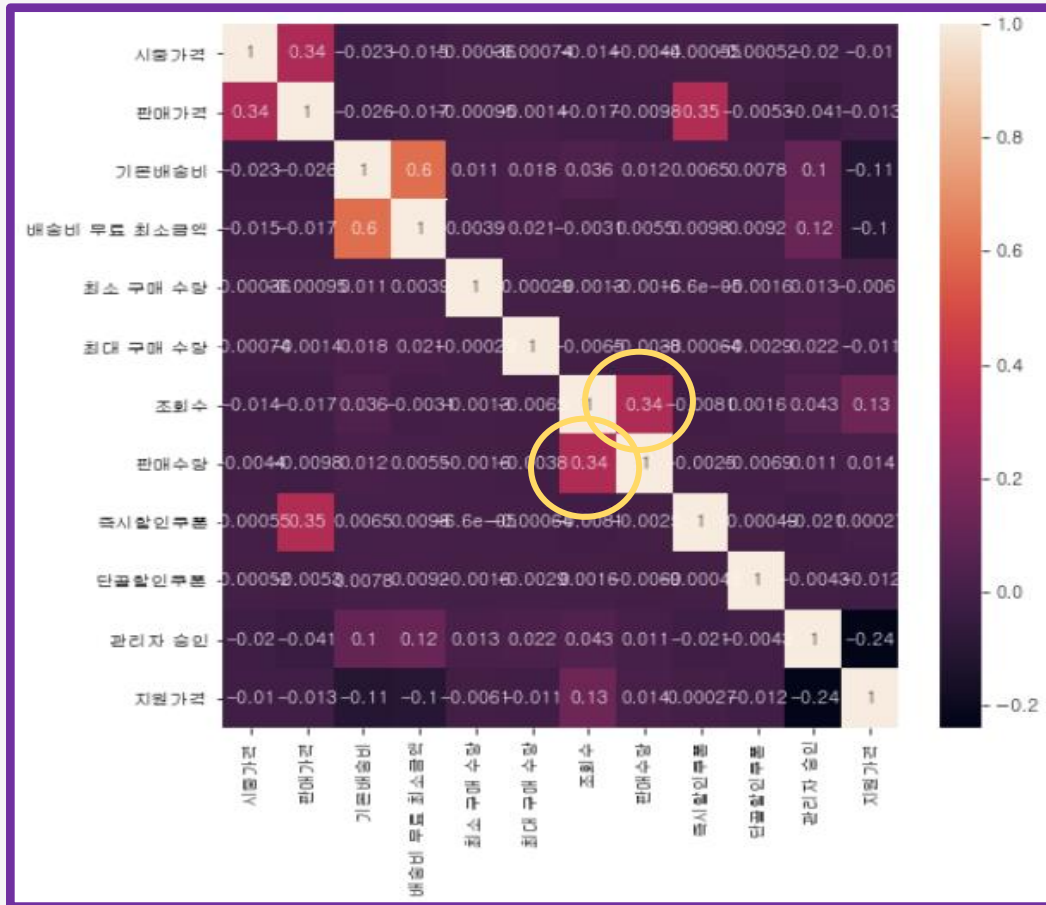
데이터 전처리

B
기본분류
7020
201030
7020
3020
3020

zz값(이벤트 상품)을 제외하여 일반적인 상황의 상품 추천

1차, 2차, 3차 분류를 모두 기본분류로 통합

상품 선호도 테이블 구성



이용자 개인 맞춤형 서비스를 위해선 사용자의 행동 양식 데이터가 필요!

-> Heatmap Chart 그래프를 통해
상품 관리 데이터 간 상관관계 비교

-> 조회수와 판매수량의 상관관계가 0.34로
두 Parameter 간 충분한 상관관계가 있음을 가정

상품 선호도 테이블 구성

	상품코드	기본분류	분류명	상품유형	상품명	제조사	판매가격	기본판매가	조회수	판매수량	업체ID	선호도
0	157092753	201030	꽃감/반건시/감말랭이	NaN	[상주전통꽃감] 전통꽃감 건시 4kg(65~70과)	상주전통꽃감	89000	0	0	0	d3f908bbea5d2d7cac866508dff9b2ffdf0088ab55d178...	22.961649
1	1578008979	201030	꽃감/반건시/감말랭이	농산물	[구릿돌농원] 건시3호 48구	구릿돌농원	55000	4000	130	0	b4f5eabe0dd4a2502f409f9710682f72bbbc79de195121...	19.849017
2	1578031964	201030	꽃감/반건시/감말랭이	농산물	[구릿돌농원] 건시4호 36구	구릿돌농원	45000	4000	223	1	b4f5eabe0dd4a2502f409f9710682f72bbbc79de195121...	18.306281
3	1578034429	201030	꽃감/반건시/감말랭이	농산물	[구릿돌농원] 건시 날개 포장 30구	구릿돌농원	40000	4000	141	0	b4f5eabe0dd4a2502f409f9710682f72bbbc79de195121...	19.623925
4	1578034546	201030	꽃감/반건시/감말랭이	농산물	[구릿돌농원] 건시 날개 포장1호	구릿돌농원	30000	4000	136	0	b4f5eabe0dd4a2502f409f9710682f72bbbc79de195121...	19.725603
...
3124	1604998092	80	김치/장류/참기름	가공식품	바른간장 1.8L	바른된장	24000	3000	186	1	926268d8bc66615558bacbd42caa33ac24a0d0c7d032fd...	18.937150
3125	1606396480	80	김치/장류/참기름	가공식품	두레원무염청국장 1kg	두레원식품	13000	3500	105	0	a279427a10453ba64bc3ff604c06d63769647b4df7a5d5...	20.380307
3126	1606924853	80	김치/장류/참기름	가공식품	두레원무염청국장 200g	두레원식품	3200	3500	194	45	a279427a10453ba64bc3ff604c06d63769647b4df7a5d5...	27.140965
3127	1606196135	b040	참기름	공산품	[풍기인견 아이싱] 사각 끈 매밀베게	풍기인견 아이싱	23000	0	3	0	1e5162a3dd797424426992b764688a7e7759b85a2597d9...	22.878855
3128	1606296202	e0	이전_농가채형	NaN	★제550수요특가★ 골고구마 상 10킬로 영주 능가 직접배송! [하늘기둥농장]	하늘기둥농장	43000	4000	3241	26	860422ba909816eacec97a4f09042e53fe1b2b1af39df8...	25.381132

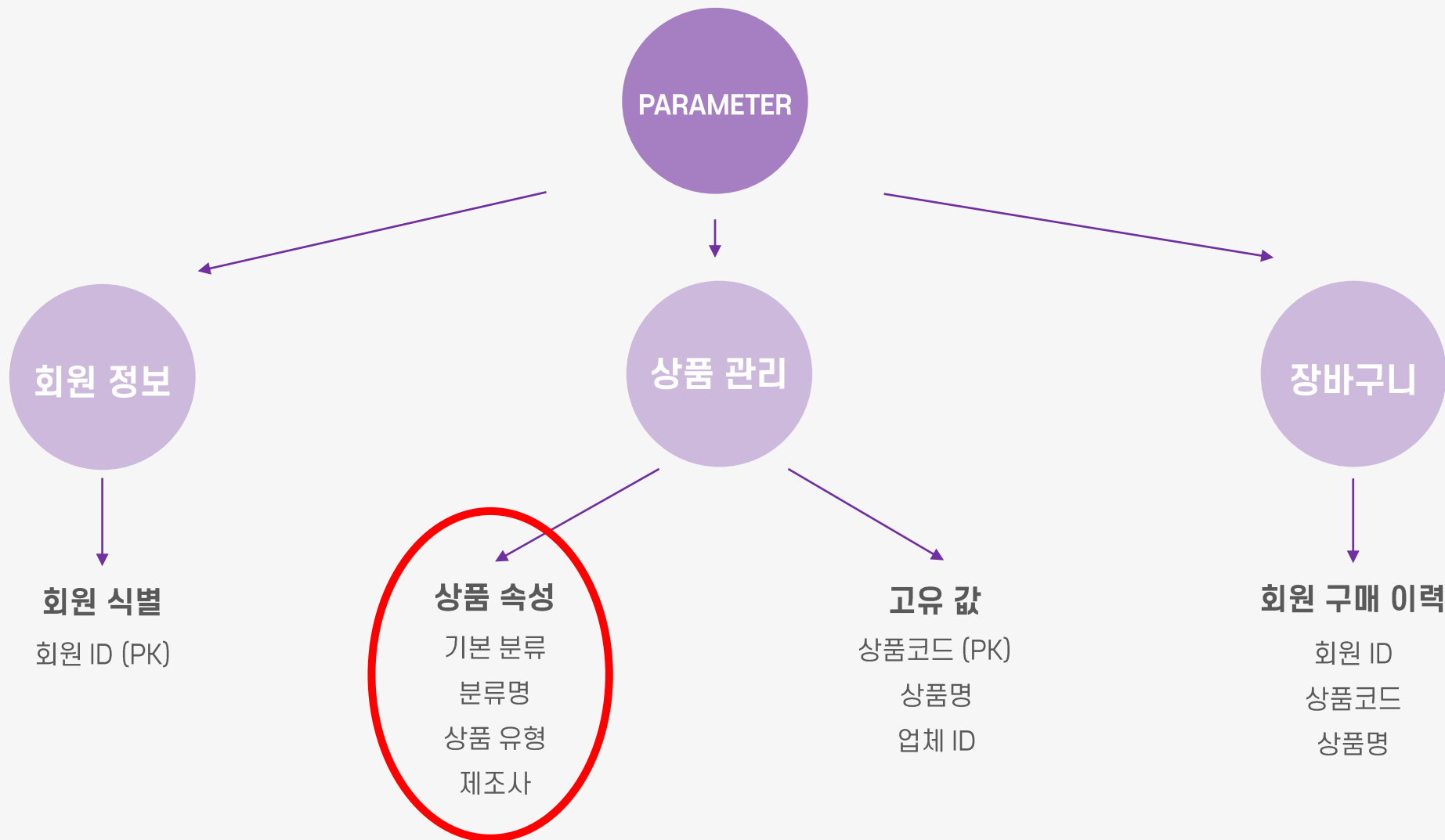
3129 rows x 12 columns

-> 이후 조회수와 판매량을 이용하여 Weighted Rating 연산 과정을 수행하였고, 선호도 데이터를 얻을 수 있었다

$$\text{weighted rating (WR)} = (v \div (v+m)) \times R + (m \div (v+m)) \times C$$

- R: 개별 상품의 판매 수량 (Rating)
- v: 개별 상품에 대한 조회수 (Votes)
- m: 상위 500위 내에 포함되기 위한 최소 조회수 (minimum votes)
- C: 전체 상품의 평균 판매 수량

PARAMETER 설정



콘텐츠 기반 필터링 모델

▼ 입력 받은 상품명을 기반으로, 상품과 유사한 다른 상품을 추천하는 함수

```
✓ 0초 ▶ def get_recommend_merchandise_list(df, merc_name, top):  
  
    # 특정 상품과 비슷한 상품을 추천해야 하므로, '특정 상품' 정보 추출  
    target_merc_index = df[df["상품명"] == merc_name].index.values  
  
    # 코사인 유사도 중 비슷한 코사인 유사도를 가진 정보 추출  
    sim_index = category_c_sim[target_merc_index, :top].reshape(-1)  
    # 자기 자신을 제외  
    sim_index = sim_index[sim_index != target_merc_index]  
  
    # data frame으로 만들고, '선호도'로 정렬한 뒤 return  
    result = df.iloc[sim_index].sort_values('선호도', ascending=False)[:top]  
    return result
```

콘텐츠 기반 필터링 모델

▶ 입력 받은 상품명을 기반으로, 상품과 유사한 다른 상품을 추천하는 함수

```
✓ 0초 ▶ def get_recommend_merchandise_list(df, merc_name, top):  
  
    # 특정 상품과 비슷한 상품을 추천해야 하므로, '특정 상품' 정보 추출  
    target_merc_index = df[df["상품명"] == merc_name].index.values  
  
    # 유사한 유사도 중 비슷한 유사도 중 유사도를 가장 많은 상품  
    siml_index = df[df["similarity"] > target_merc_index].index.values  
    # 가장 유사한 상품  
    siml_index = siml_index[siml_index != target_merc_index]  
  
    # data frame으로 만들고, '유사도'를 정렬한 뒤 return  
    result = df[siml_index].sort_values("유사도", ascending=False).head  
    return result
```


콘텐츠 기반 필터링 모델

• 입력 받은 상품명을 기반으로, 상품과 유사한 다른 상품을 추천하는 함수

```
def get_recommendation_by_content(target_name, merc_index, top):  
    """  
    # 특정 상품과 유사한 상품을 추천해주는 함수로, 특정 상품 정보 추출  
    target_merc_index = merc_index[상품명] + 1  
    """  
  
    # 코사인 유사도 중 비슷한 코사인 유사도를 가진 정보 추출  
    sim_index = category_csim[target_merc_index, :top].reshape(-1)  
    # 자기 자신을 제외  
    sim_index = sim_index[sim_index != target_merc_index]  
  
    # data_frame으로 만들고, 유사도 순 정렬한 뒤 return  
    result = pd.DataFrame(index=merc_index, columns=[상품명, 'recommender'], data=  
        sim_index)
```

콘텐츠 기반 필터링 모델

• 입력 받은 상품명을 기반으로, 상품과 유사한 다른 상품을 추천하는 함수

```
def get_recommendation_name(target_name, item_index, top):  
    # 특정 상품과 유사한 상품을 추천해야 하므로, 특정 상품 정보 추출  
    target_item_index = df[df['상품명'] == target_name].index.values  
  
    # 유사한 상품을 잘 유사한 상품을 가진 상품 추출  
    sim_index = df[df['similarity'] == target_item_index].index.values  
    # 유사 상품 추출  
    sim_index = sim_index[sim_index != target_item_index]  
  
    # data frame으로 만들고, '선호도'로 정렬한 뒤 return  
    result = df.iloc[sim_index].sort_values('선호도', ascending=False)[:top]  
    return result
```

콘텐츠 기반 필터링 모델

```
✓ 0초 # 임의의 상품명  
merc_name = "[구릿뜰농원] 건시5호 24구"  
top = 5  
  
get_recommend_merchandise_list(contents, merc_name, top)  
# 분류명을 기준으로, 위 상품과 유사한 상품 5개를 추천
```

상	상품코드	기본분류	분류명	상품유형	상품명
0	157092753	201030	꽃감/반건시/감말랭이	NaN	[상주전통꽃감] 전통꽃감 건시 4kg(65~70과)
85	1600253088	201030	꽃감/반건시/감말랭이	NaN	★한가위할인판매★ 구릿뜰꽃감 반건시2호(3.2kg40개)
83	1600043466	201030	꽃감/반건시/감말랭이	농산물	★이웃사촌★ [선물용]서리골햇꽃감(1KG 20과~24과)
84	1600129353	201030	꽃감/반건시/감말랭이	농산물	프리미엄 개별건시 1호
73	1599056687	201030	꽃감/반건시/감말랭이	농산물	[상주아람꽃감]상주꽃감 오동선물세트(50g 48개)

상품명	
[상주전통꽃감] 전통꽃감 건시 4kg(65~70과)	
★한가위할인판매★ 구릿뜰꽃감 반건시2호(3.2kg40개)	
★이웃사촌★ [선물용]서리골햇꽃감(1KG 20과~24과)	
프리미엄 개별건시 1호	
[상주아람꽃감]상주꽃감 오동선물세트(50g 48개)	

필터링 모델을 사용한 결과

Input 값과 '상품명'의 코사인 유사도가 비슷한 값들이 추천됨

콘텐츠 기반 필터링 모델

```
# 임의의 상품명
merc_name = "[구릿뜰농원] 건시5호 24구"
top = 5
```

```
[17] # 임의의 상품명
merc_name = "[구릿뜰농원] 건시5호 24구"
top = 5

get_recommend_merchandise_list(contents, merc_name, top)
# 제조사를 기준으로, 위 상품과 유사한 상품 5개를 추천
```

	상품코드	기본분류	분류명	상품유형	상품명	제조사	판매가격	기본배송비	조회수	판매수량	업체 ID	선호도
0	157092753	201030	꽃감/반건시/감말랭이	NaN	[상주전통꽃감] 전통꽃감 건시 4kg(65~70과)	상주전통꽃감	89000	0	0	0	d3f908bbea5d2d7cac866508dff9b2ffd0088ab55d178...	22.961649
85	1600253088	201030	꽃감/반건시/감말랭이	NaN	★한가위할인판매★ 구릿뜰꽃감 반건시2호(3.2kg40개)	구릿뜰농원	50000	3000	130	0	b4f5eabe0dd4a2502f409f9710682f72bbbc79de195121...	19.849017
83	1600043466	201030	꽃감/반건시/감말랭이	농산물	★이웃사촌★ [선물용]서리골햇꽃감(1KG 20과~24과)	내서 서리골마을	25500	4000	577	13	5f90d34c68fb01f90df340c684d12fdb25346c06af0853...	18.873547
84	1600129353	201030	꽃감/반건시/감말랭이	농산물	프리미엄 개별건시 1호	상주한시꽃감	50000	0	182	0	867288f154a3d1d3e3aa90236e341a3f82b7f69050d9fd...	18.828098
73	1599056687	201030	꽃감/반건시/감말랭이	농산물	[상주아람꽃감]상주꽃감 오동선물세트(50g 48개)	상주아람꽃감농장	90000	0	248	0	a37d03f03447c79d311ac6caa770bbdbec75fc55d04f7b...	17.674287

다양한 속성에 따라 다양한 추천이 가능하다는 강점을 가지고 있음

필터링 모델을 사용한 결과

Input 값과 '상품명'의 코사인 유사도

```
[27] # 임의의 상품명
merc_name = "[구릿뜰농원] 건시5호 24구"
top = 5
```

```
get_recommend_merchandise_list(contents, merc_name, top)
# 업체 ID를 기준으로, 위 상품과 유사한 상품 5개를 추천
```

	상품코드	기본분류	분류명	상품유형	상품명	제조사	판매가격	기본배송비	조회수	판매수량	업체 ID	선호도
33	1585313479	201030	꽃감/반건시/감말랭이	농산물	상주구릿뜰꽃감 건시5호	구릿뜰농원	28000	4000	6	0	b4f5eabe0dd4a2502f409f9710682f72bbbc79de195121...	22.796655
1	1578008979	201030	꽃감/반건시/감말랭이	농산물	[구릿뜰농원] 건시3호 48구	구릿뜰농원	55000	4000	130	0	b4f5eabe0dd4a2502f409f9710682f72bbbc79de195121...	19.849017
4	1578034546	201030	꽃감/반건시/감말랭이	농산물	[구릿뜰농원] 건시 날개 포장1호	구릿뜰농원	30000	4000	136	0	b4f5eabe0dd4a2502f409f9710682f72bbbc79de195121...	19.725603
3	1578034429	201030	꽃감/반건시/감말랭이	농산물	[구릿뜰농원] 건시 날개 포장 30구	구릿뜰농원	40000	4000	141	0	b4f5eabe0dd4a2502f409f9710682f72bbbc79de195121...	19.623925
2	1578031964	201030	꽃감/반건시/감말랭이	농산물	[구릿뜰농원] 건시4호 36구	구릿뜰농원	45000	4000	223	1	b4f5eabe0dd4a2502f409f9710682f72bbbc79de195121...	18.306281

아이템 기반 필터링 모델

아이템 기반 필터링 모델에 사용할 PARAMETER 설정

회원 정보 테이블에서 '회원번호' 추출, 상품 관리 테이블에서 '상품명' 추출, 연산을 통해 얻은 '선호도' 추출

회원번호	1	60251	60603	60729	60846	61002	61062	61184	61484	61677	61772	61775	61891	61926	62191	62200	62241
상품명																	
"힘내세요, 우리나라" 기세음[곰뽕이(꽃뽕이)진액/액기스]	16.495395	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
(경북청년농부들) 민서네버섯농장간식라이스표고버섯 150g	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
(문경약동돼지) 문경약동돼지 LA식갈비(냉동) 1kg	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

아이템 기반 PIVOT 테이블 구성

Index = 상품명

Column = 회원번호

Data = 선호도

이를 통해서 회원이 구매한 상품과 선호도를 한 테이블 내에서 확인 가능

아이템 기반 필터링 모델

추천 함수 생성

입력한 상품명과 유사한 선호도를 가지는 상품을 추천하는 함수 생성

```
[ ] # 입력한 상품(상품명을 입력)과 유사한(코사인 유사도) 6개의 상품을 추천하는 기능의 함수
def get_item_based_collabor(merc_name):
    return item_based_collabor[merc_name].sort_values(ascending=False)[:6]
```

임의의 상품명 입력

get_item_based_collabor('피크닉사과 가정용 홈과 10kg')

상품명

답운농장 큰며느리가 만든 현미누룽지, 흑임자[검정깨]누룽지,새싹보리누룽지,홍삼누룽지,천년초누룽지	0.169031
[하늘기둥농장] 간편 소포장 청국장 2킬로(200그램*10)	0.169031
명품안동소주 담금주 3600ml 30도	0.169031
★미웃사촌★[달산마을]햇 알호두(30알)380g~400g	0.119523
[윤푸드] 미로우니 우리쌀 연근돈까스 650g (130g*5개)	0.119523
[대우농장] 완숙 토마토 5kg (특품)	0.119523

코사인 유사도를 통해 측정 후 추천

선호도가 높은 상품만을 추천하는 것이 아닌,
사용자의 행동양식을 기반으로 한 추천이라는 점이 강점

결과적으로, 다양한 상품을 이용자 개인 맞춤형으로 추천

구매이력 기반 잠재요인 필터링 모델

상품코드	1577870016	1577933738	1577947433	1577948991	1577950662	1577971381	1578032329	157804345
회원번호								
1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
60251	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
60596	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
60603	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
60729	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
...
169279	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
169280	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
169282	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
169286	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
169295	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

8988 rows x 1590 columns

Pivot 테이블 생성

구매이력 파악을 위한
고객-상품 선호도 Pivot 테이블 생성

상품코드	1577870016	1577933738	1577947433	1577948991	1577950662	1577971381	1578032329	157804345
0	-0.273593	-0.273593	-0.273593	-0.273593	-0.273593	-0.273593	-0.273593	-0.273593
1	-0.071004	-0.071004	-0.071004	-0.071004	-0.071004	-0.071004	-0.071004	-0.071004
2	-0.088722	-0.088722	-0.088722	-0.088722	-0.088722	-0.088722	-0.088722	-0.088722
3	-0.023095	-0.023095	-0.023095	-0.023095	-0.023095	-0.023095	-0.023095	-0.023095
4	-0.083072	-0.083072	-0.083072	-0.083072	-0.083072	-0.083072	-0.083072	-0.083072
...
8983	-0.017102	-0.017102	-0.017102	-0.017102	-0.017102	-0.017102	-0.017102	-0.017102
8984	-0.018033	-0.018033	-0.018033	-0.018033	-0.018033	-0.018033	-0.018033	-0.018033
8985	-0.010583	-0.010583	-0.010583	-0.010583	-0.010583	-0.010583	-0.010583	-0.010583
8986	-0.009509	-0.009509	-0.009509	-0.009509	-0.009509	-0.009509	-0.009509	-0.009509
8987	-0.020094	-0.020094	-0.020094	-0.020094	-0.020094	-0.020094	-0.020094	-0.020094

8988 rows x 1590 columns

선호도 연산

m개의 상품에 대한 n명의 평균 선호도를 구한 뒤,
고객-상품 선호도 테이블에서 뺌

$$\begin{matrix} M \\ m \times n \end{matrix} = \begin{matrix} U \\ m \times m \end{matrix} \begin{matrix} \Sigma \\ m \times n \end{matrix} \begin{matrix} V^* \\ n \times n \end{matrix}$$

잠재요인 추출

Truncated SVD를 통해
행렬분해를 거치고 잠재요인 추출

구매이력 기반 잠재요인 필터링 모델

```
def recommend_merchandise(df_svd_preds, customer_id, ori_merc_df, ori_ratings_df, num_recommendations):  
    # 최종적으로 만든 pred_df에서 고객 index에 따라 상품 데이터 정렬하기 -> 상품 선호도가 높은 순으로 정렬됨  
    sorted_customer_pred = df_svd_preds.loc[df_svd_preds["회원번호"] == customer_id].set_index("회원번호").transpose().sort_values(by=customer_id, ascending=False).transpose()  
  
    # 원본 선호도 데이터에서 회원번호에 해당하는 데이터 추출  
    customer_data = df_ratings[df_ratings["회원번호"] == customer_id]  
  
    # 추출한 회원 데이터와 원본 상품 데이터를 합침  
    customer_history = customer_data.merge(df_merchandise, on='상품코드').sort_values(['선호도'], ascending=False)  
  
    # 원본 상품 데이터에서 고객의 구매한 상품을 제외한 데이터를 추출  
    recommendations = df_merchandise[df_merchandise['상품코드'] != sorted_customer_pred['상품코드']]  
  
    # 고객의 선호도가 높은 순으로 정렬한 데이터로, 위의 recommendations 데이터를 합침  
    recommendations = recommendations.merge(sorted_customer_pred, on='회원번호', left_index=True, right_index=False)  
  
    # 고객의 선호도가 높은 순으로 정렬한 데이터로, 위의 recommendations 데이터를 합침  
    recommendations = recommendations.merge(customer_data, on='상품코드', left_index=True, right_index=False, ascending=False, suffixes=(None, '_customer_data'))  
  
    return customer_history, recommendations
```


구매이력 기반 잠재요인 필터링 모델

```
def recommend merchandise(df_merchandise, customer_id, sort_order, num_recommendations):  
    # 상품코드별로 판매 수를 세고, 고객 선호도에 따라 상품 코드를 정렬한다. 선호도가 높은 순으로 정렬  
    sorted_customer_pred = df_merchandise.groupby('상품코드')['customer_id'].transform(lambda x: x.sort_values(ascending=False).transform(lambda x: x.head(1))))  
    # 판매 수를 세고, 고객 선호도에 따라 상품 코드를 정렬한다. 선호도가 높은 순으로 정렬  
    sorted_customer_pred = df_merchandise.groupby('상품코드')['customer_id'].transform(lambda x: x.sort_values(ascending=False).transform(lambda x: x.head(1))))  
    # 추천할 상품 코드를 선택한다. 선호도가 높은 순으로 정렬된 상품 코드를 선택  
    sorted_customer_pred = df_merchandise.groupby('상품코드')['customer_id'].transform(lambda x: x.sort_values(ascending=False).transform(lambda x: x.head(1))))  
    # 원본 상품 데이터에서 고객이 구매한 상품을 제외한 데이터를 추출  
    recommendations = df_merchandise[~df_merchandise['상품코드'].isin(customer_history['상품코드'])]  
    # 고객의 선호도가 높은 순으로 정렬된 데이터와, 위의 recommendation 데이터를 합침  
    recommendations = recommendations.merge(pd.DataFrame(sorted_customer_pred.transpose()).reset_index(), on='상품코드')  
    # 고객의 선호도가 높은 순으로 정렬된 데이터와, 위의 recommendation 데이터를 합침  
    recommendations = recommendations.rename(columns={customer_id: "추천 예상"}).sort_values(by="추천 예상", ascending=False).head(num_recommendations)  
    return customer_history, recommendations
```

구매이력 기반 잠재요인 필터링 모델

already_rated

회원 번호	상품코드	선호도	상품명	배송일	배송 완료 일	
0	60729	1598427460	57.537618	[아름드리사과] 2020 햇사과 가정용 홀로사과 5kg 소과 20-23과내	2020-08-31 00:00:00	2020-09-01 00:00:00
37	60729	1598427460	57.537618	[아름드리사과] 2020 햇사과 가정용 홀로사과 5kg 소과 20-23과내	2020-09-08 00:00:00	2020-09-09 00:00:00
53	60729	1598427460	57.537618	[아름드리사과] 2020 햇사과 가정용 홀집 빨간사과 5kg 소과 20-23과내	2020-10-05 00:00:00	2020-10-06 00:00:00
53	60729	1598427460	57.537618	[아름드리사과] 2020 햇사과 가정용 홀집 빨간사과 5kg 소과 20-23과내	2020-09-05 00:00:00	2020-09-06 00:00:00

고객 구매이력 출력

[] recommendations

장바구니번호	상품코드	상품명	추천 예상
10205	2649252	1.592374e+09 20년 속이 꽉찬 유기농 양파 (중/대) 3kg, 5kg, 10kg	0.100504
10153	2648439	1.592374e+09 20년 속이 꽉찬 유기농 양파 (중/대) 3kg, 5kg, 10kg	0.100504
10146	2646698	1.592374e+09 20년 속이 꽉찬 유기농 양파 (중/대) 3kg, 5kg, 10kg	0.100504
10147	2647097	1.592374e+09 20년 속이 꽉찬 유기농 양파 (중/대) 3kg, 5kg, 10kg	0.100504
10148	2648245	1.592374e+09 20년 속이 꽉찬 유기농 양파 (중/대) 3kg, 5kg, 10kg	0.100504

구매 이력을 기반으로 한 추천 리스트

구매이력 기반 잠재요인 필터링 모델

추출한 잠재요인을 통해 측정할 수 없는 선호도마저 예측 가능

already_rated

회원 번호	상품코드	선택도	상품명	배송일	배송원 요배정 일
0	60729	1598427460	57.537618	[아름드리사과] 2020 햇사과 가정용 중로사과 5kg 소과 20-23과내	2020-08-31 2020-09-01 00:00:00 00:00:00
37	60729	1598427460	57.537618	[아름드리사과] 2020 햇사과 가정용 중로사과 5kg 소과 20-23과내	2020-08-08 2020-09-09 00:00:00 00:00:00
53	60729	1598427460	57.537618	[아름드리사과] 2020 햇사과 가정용 중로사과 5kg 소과 20-23과내	2020-10-05 2020-10-06 00:00:00 00:00:00
53	60729	1598427460	57.537618	[아름드리사과] 2020 햇사과 가정용 중로사과 5kg 소과 20-23과내	2020-09-25 2020-09-26 00:00:00 00:00:00

고객 구매이력 출력

[] recommendations

장바구니번호	상품코드	상품명	추천 배상
10105	2649212	1.592374e+09 20년 속이 폭찬 유기농 양파 (중/대) 3kg, 5kg, 10kg	0.100504
10145	2648245	1.592374e+09 20년 속이 폭찬 유기농 양파 (중/대) 3kg, 5kg, 10kg	0.100504
10146	2646698	1.592374e+09 20년 속이 폭찬 유기농 양파 (중/대) 3kg, 5kg, 10kg	0.100504
10147	2647097	1.592374e+09 20년 속이 폭찬 유기농 양파 (중/대) 3kg, 5kg, 10kg	0.100504
10148	2648245	1.592374e+09 20년 속이 폭찬 유기농 양파 (중/대) 3kg, 5kg, 10kg	0.100504

구매 이력을 기반으로 한 추천 리스트

하이브리드 추천 시스템 FLOW

콘텐츠 기반
필터링

속성 외의 다른 것들을 고려하지 않음

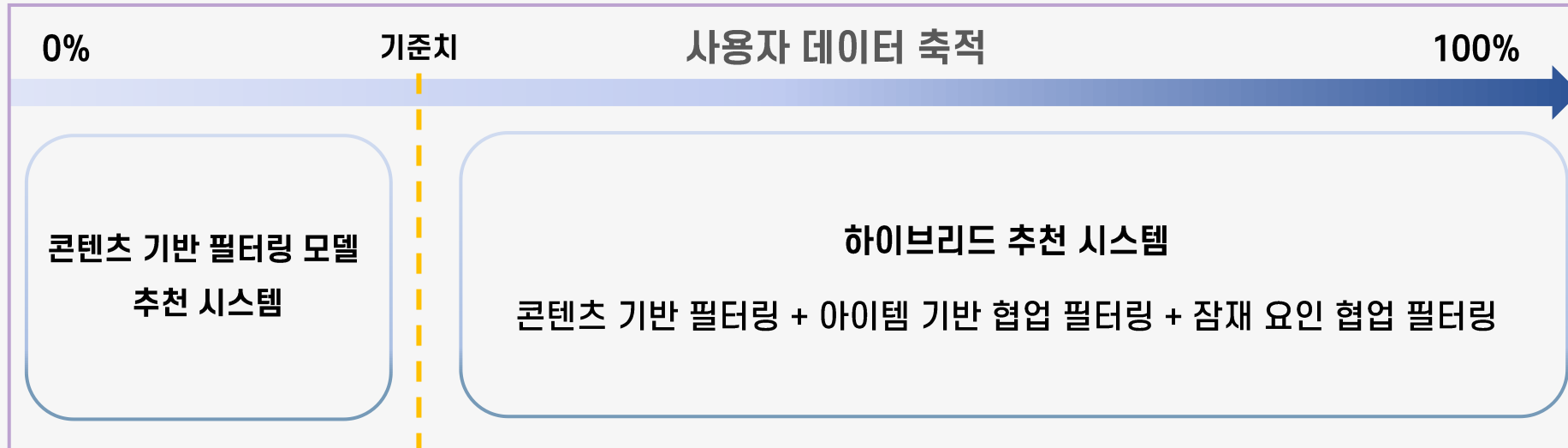
협업
필터링

데이터가 충분하지 않은 신규 회원들에게 적용 어려움



두 필터링의 장점을 살리고 단점을 보완할 수 있는 하이브리드 추천 시스템 고안

하이브리드 추천 시스템 FLOW



사용자의 데이터가 축적이 되기 전(콜드 스타트)에는
콘텐츠 기반 필터링 모델 추천 시스템 사용하여 보완 가능

데이터가 축적되면 잠재 요인 협업, 아이템 기반 협업 등
다양한 필터링 모델 사용

공공 데이터를 활용한 확장 방향



지도 API를 이용한 발전 방향

지리적 데이터를 이용해 주변 업체를 추천하는 서비스를 기반으로
픽업 서비스 등 다양한 서비스 구축

- 분야별 세부실천계획 -

⑥ 소득 걱정없는 농산어촌 경북

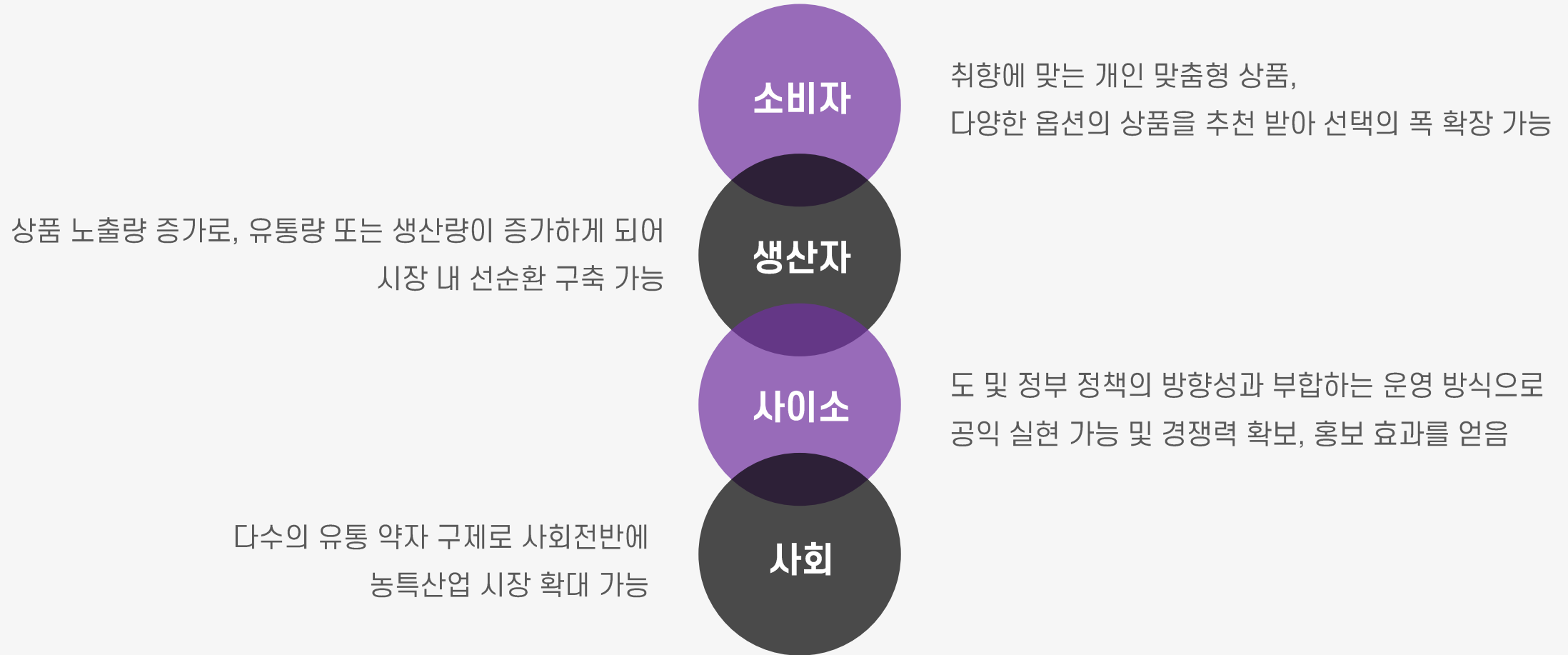
1. 경북 농식품 유통전담기관 설립
2. 청년농부 영농정착 지원 및 경북형 농업복지모델
3. 똑똑한 농업, 스마트팜 혁신밸리 구축
4. 미래 농업인재 육성, 청년 창농의 메카 경북
5. 농업 6차산업 전진기지 구축
6. 경북형 두레공동체 육성
7. 청담농업 기술개발, 지식기반 스마트농업 실현
8. 2030세대 귀농어업인 안정정착 프로그램
9. 고부가 농식품산업 및 수출농업 육성
10. 경북형 골드시드 프로젝트
11. 미래농업 블루오션, 곤충산업 활성화
12. 약취없는 청정클린 축산업 육성
13. 고소득 친환경 스마트 축산 확산
14. 임산물 6차 산업화 단지 조성
15. 첨단 가공산업, 살맛나는 어촌 만들기

농업 시장 확대를 통한 청년 리더 및 활발한 농업 구축

-> 추천 알고리즘 시스템은
시장 확대, 소비자의 증가 등의 효과를 낼 수 있음

-> 시장 확대와 함께
다양한 일자리와 청년, 귀농인들의 증대 효과 기대

서비스 가치



감사합니다