

Acorn Final Project



영양 불균형 문제 해결을 위한 CNN과 ML기반

반찬 예측 및 저녁 메뉴 추천 시스템

2023.07.12

i!i Acorn Academy

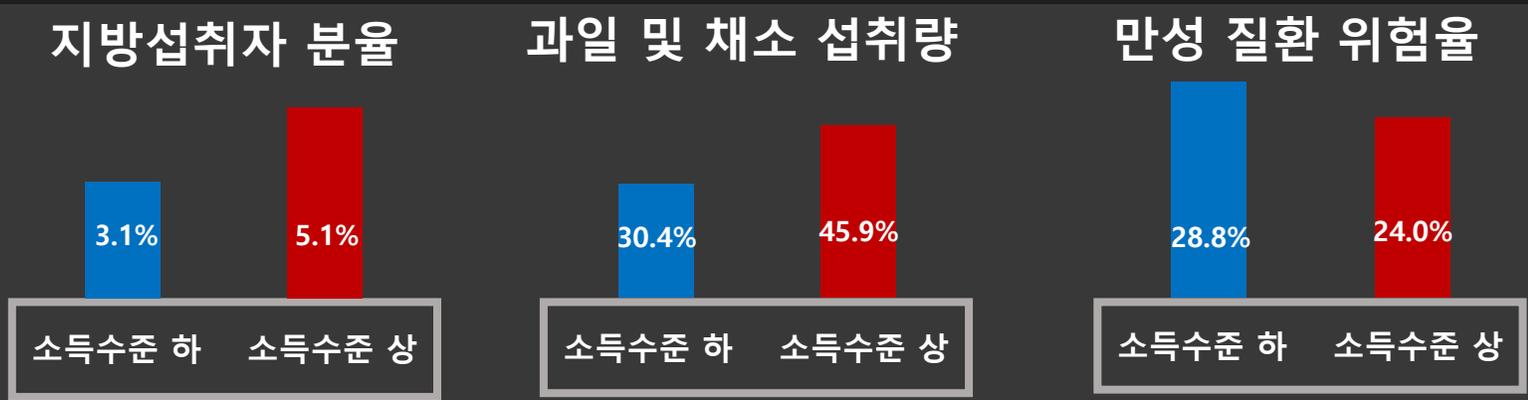
파일



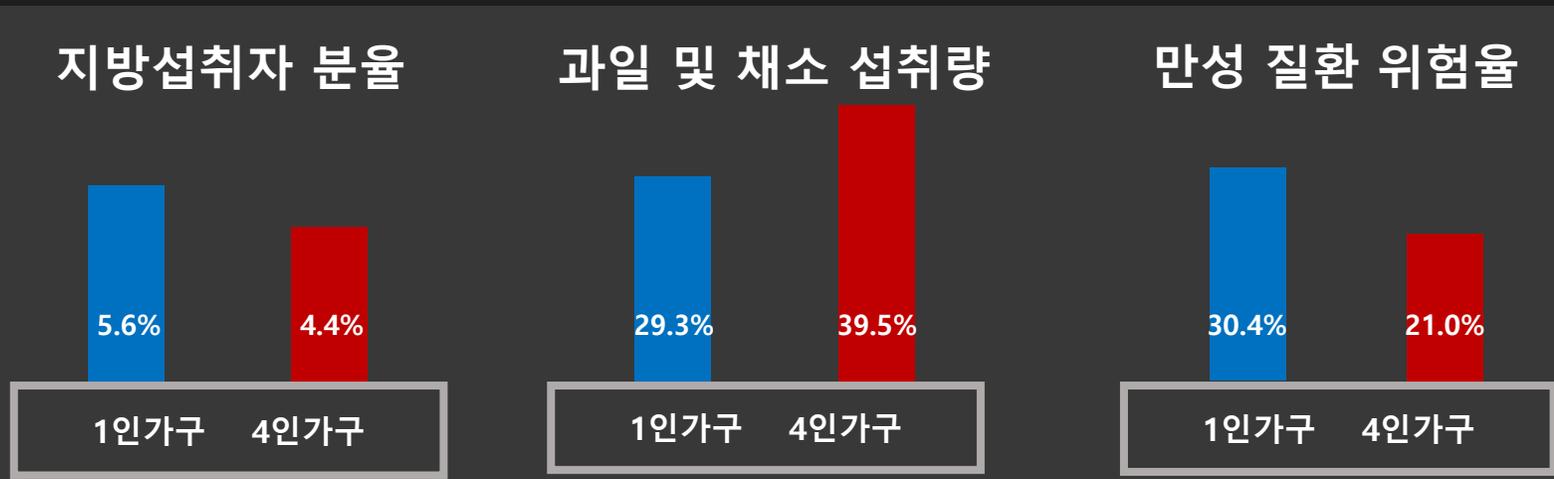
- {x} ..
- sample_data
 - 가구원수별 영양섭취부족자 비율
 - 가구소득별 고혈압 유병률
 - 가구원수별 과일채소 1일 500g ..
 - 가구원수별 에너지/지방 과잉섭 ..
 - 가구원수별 건강식생활 실천율

+ 코드 + 텍스트

```
1 plt.show("소득 수준에 따른 영양 불균형 현황")
```



```
1 plt.show("가구 인원에 따른 영양 불균형 현황")
```





+ 코드 + 텍스트

```
1 print("기존 서비스와의 차별성")
```

	음식 이미지 분류	반찬 예측	영양소 분석	메뉴 추천 서비스
	O	O	X	X
	O	O	O	X
	O	X	O	X
	O	X	O	O
 GGULKKUK	O	O	O	O



+ 코드 + 텍스트

1 Final("프로젝트명")

CNN과 ML기반 반찬 예측 및 저녁 메뉴 추천 시스템

1 Final("프로젝트 기간")

2023.06.20 ~ 2023.07.12

1 Final("멤버")

["한태성(팀장)", "김도연", "김정숙", "김호탁", "이준규", "정구열", "장상윤"]

1 Final("개발도구")

["Python", "Selenium", "Numpy", "Pandas", "Matplotlib", "Tensorflow", "Keras", "DBever", "AWS RDS", "Lucidchart", "ChatGPT"]

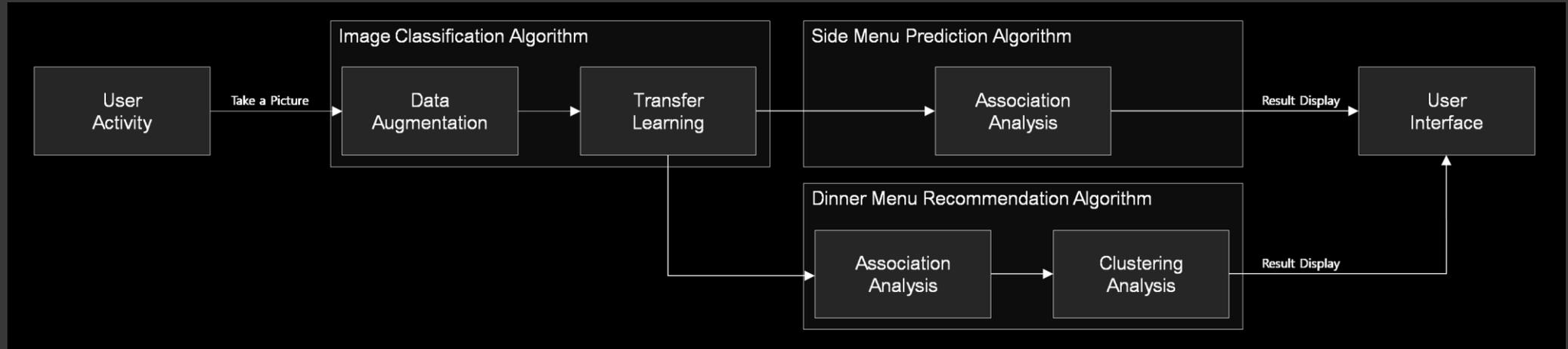
1 Final("사용 언어 및 개발 환경")

["Python 3.9.13", "Mysql Ver 15.1 Distrib 19.11 – MariaDB", "Windows 10Pro"]



+ 코드 + 텍스트

1 Final("Project Architecture")



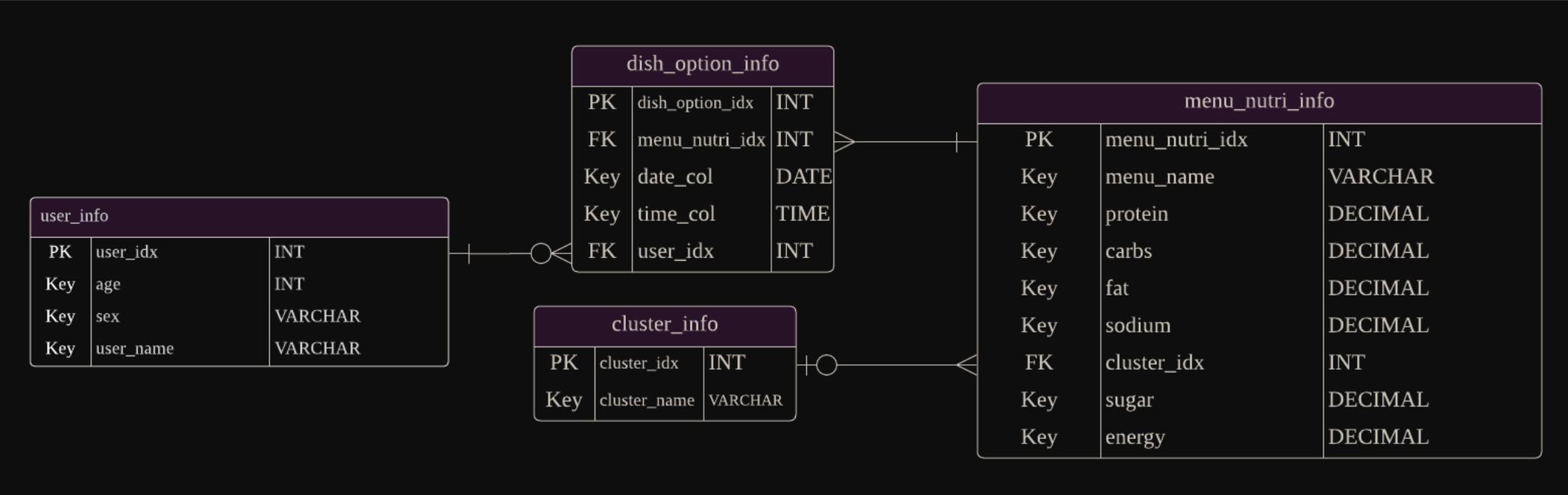
1 Final("멤버별 역할")

- ☞ "한태*" : 팀장, 총괄
- "김도*" : 프론트엔드, 기능구현
- "김정*" : 반찬예측알고리즘, 연관분석
- "김호*" : 저녁추천알고리즘, 군집분석
- "이준*" : 저녁추천알고리즘, Outlier 처리 알고리즘
- "장상*" : 백엔드, DB설계
- "정구*" : 프론트엔드, 화면설계



+ 코드 + 텍스트

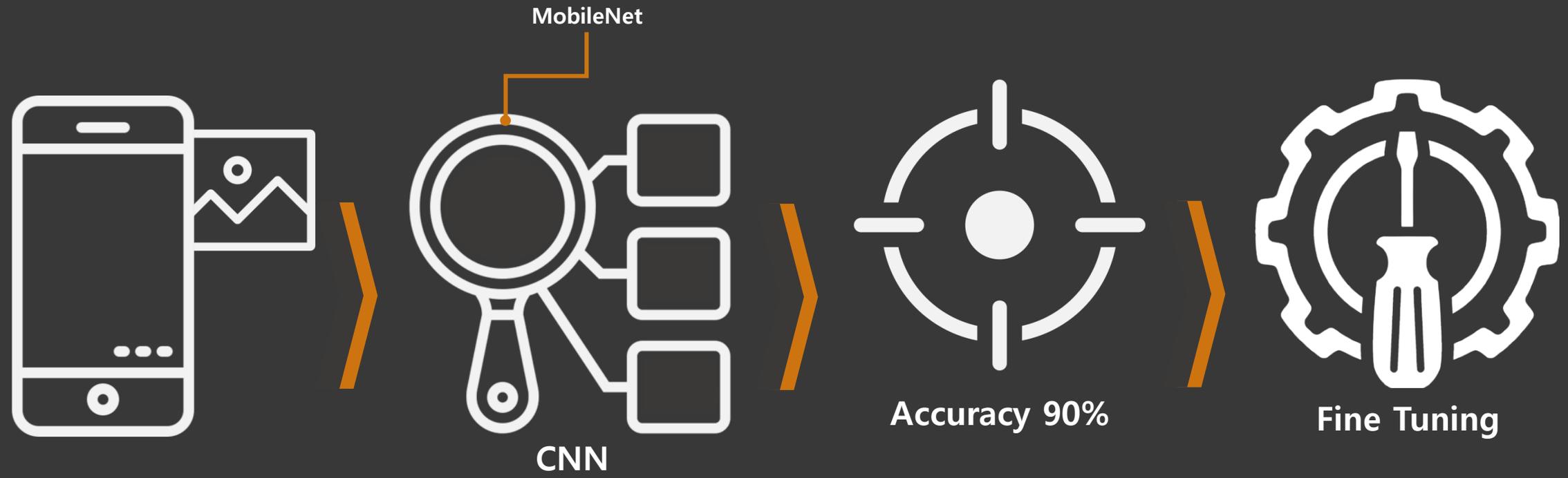
1 Final("ERD")





+ 코드 + 텍스트

1 이미지 분류 알고리즘





+ 코드 + 텍스트

1 메뉴 설정

메뉴 개수



원본 : 우리학교급식식단(식품안전나라)

1차전처리

- 오타
- 과도하게 세분화된 메뉴명
- 창의적인 메뉴명

2차 전처리

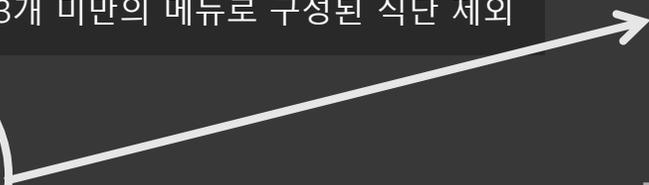
- 식품영양성분DB의 메뉴명과 매칭
- 전처리 후, 3개 미만의 메뉴로 구성된 식단 제외

메뉴 개수



비슷한 계열의 음식을 하나의 큰 카테고리로 압축
 Ex) 생선구이(고등어구이, 조기구이, 갈치구이, 임면수구이, ...)

기본 밥류와 반찬에 가까운 메뉴 제거
 Ex) 차조밥, 콩밥, 소세지구이, 계란후라이, 계란장조림, ...

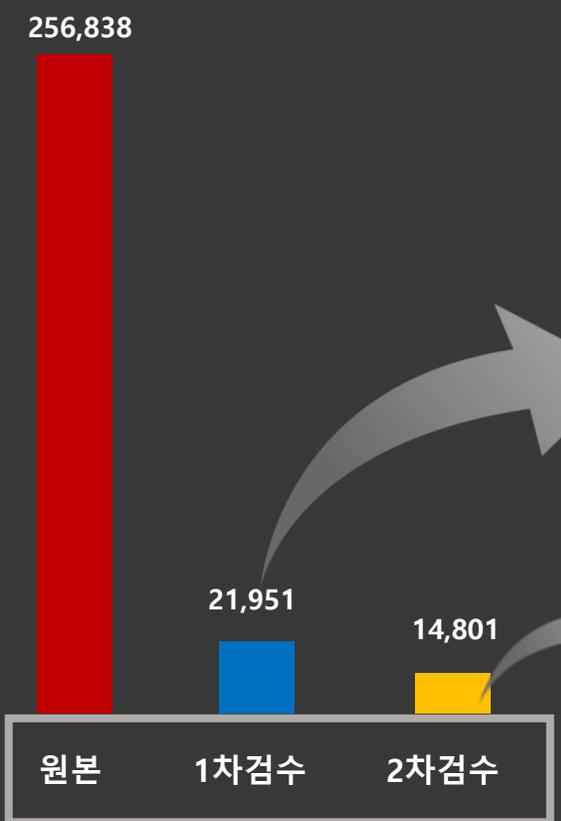




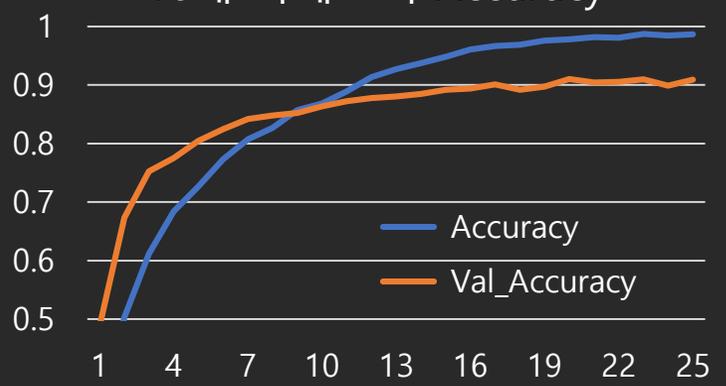
+ 코드 + 텍스트

1 이미지 수집

이미지 개수

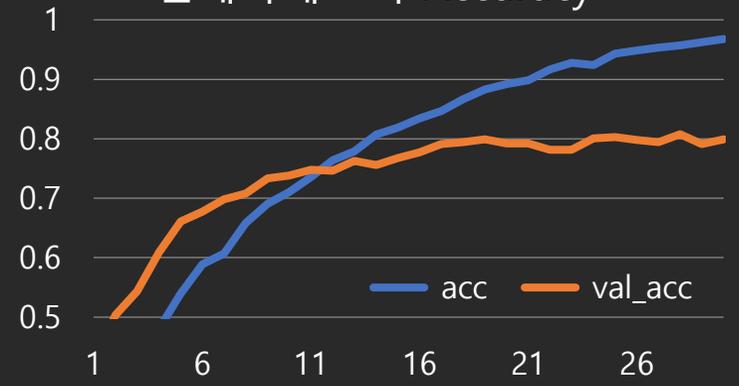


10개 카테고리 Accuracy



- 할당된 10개 카테고리별 CNN모델 평가
⇒ 정확도 약 90 %

전체카테고리 Accuracy



- 카테고리 통합 후, 전체 카테고리 CNN모델 평가
⇒ 정확도 약 80%로 감소

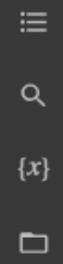
카테고리 축소 (72개→59개)



이미지 재검수



Accuracy 변동없음



+ 코드 + 텍스트



1 오답노트

1,49.0,9



2,36.0,12



6,15.0,37



8,16.0,19



10,55.0,25



12,25.0,45



23,55.0,20



28,42.0,18



파일

코드 텍스트



- Testing
- Training
 - 갈비찜

- 135degrees_00001
- 135degrees_00002
- 135degrees_00003
- 135degrees_00004
- 135degrees_00005
- 135degrees_00006
- 135degrees_00007
- 135degrees_00008
- 135degrees_00009
- 135degrees_00010
- 135degrees_00011
- 135degrees_00012
- 135degrees_00013
- 135degrees_00014
- 135degrees_00015

1 matrix = cv2.getRotationMatrix2D

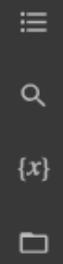
```
[ '부대찌개', '비빔국수', '탕수육', '떡국', '해물찜', '삼계탕', '생선조림', '김밥', ..., '제육볶음' ]
train label: 부대찌개 => 3488
```



```
train label: 순두부찌개 => 3536
```

데이터 개수

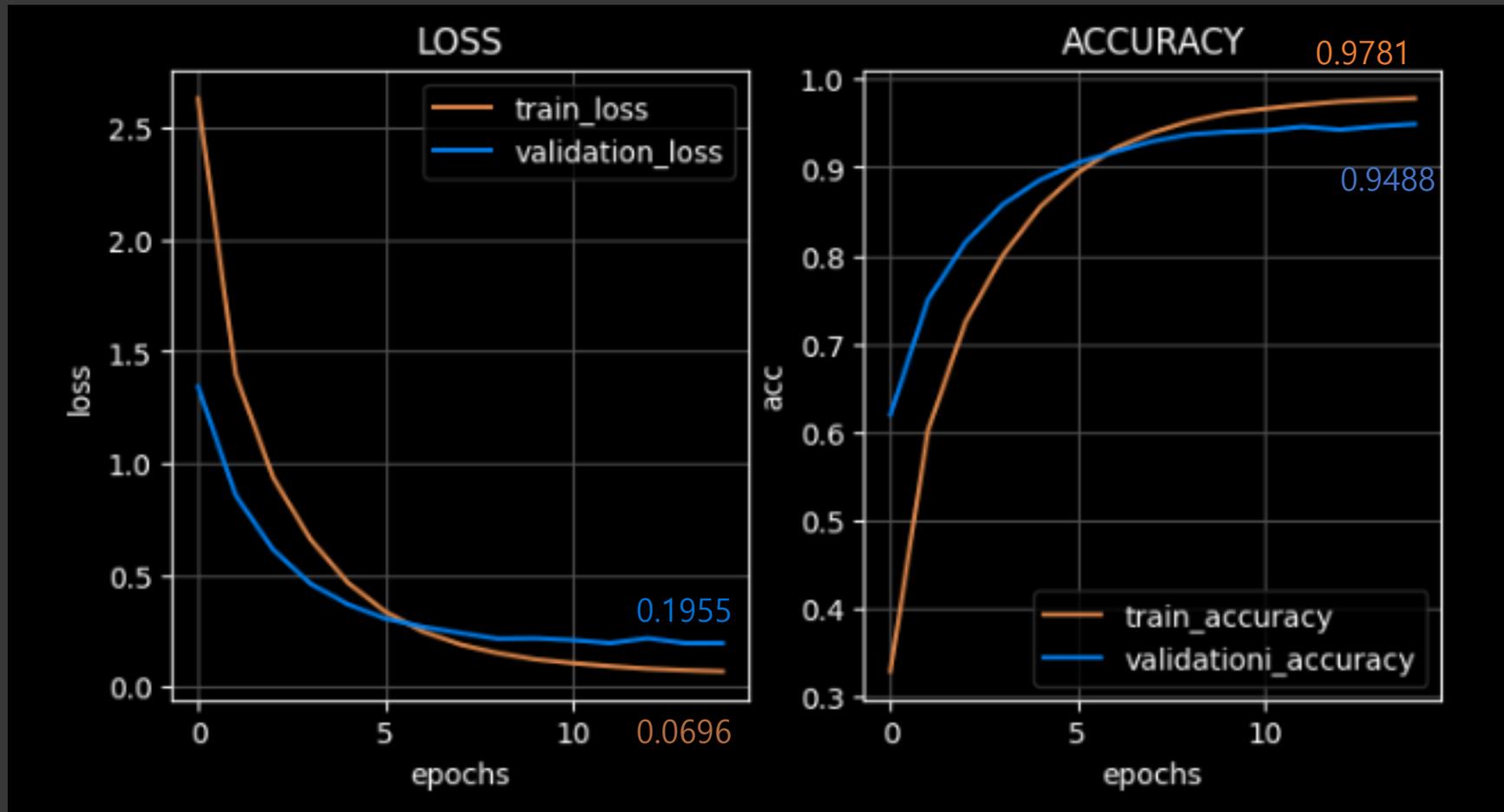
증강 전 증강 후

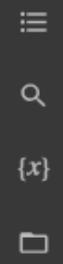


+ 코드 + 텍스트

1 Food_Image_Classification_Model

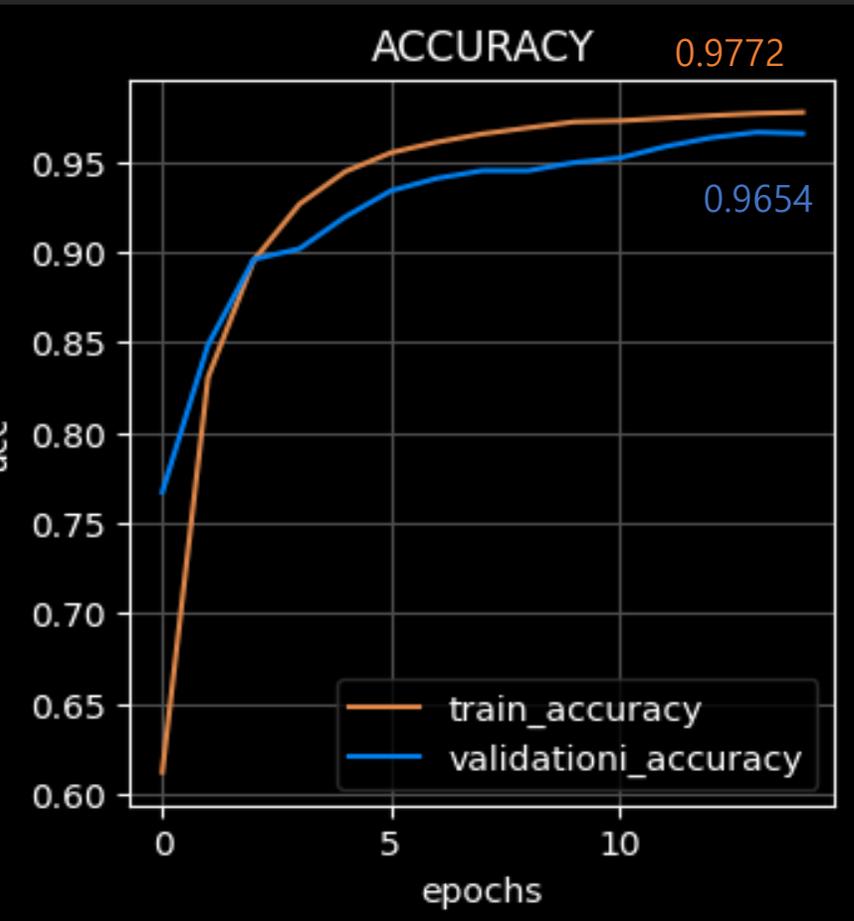
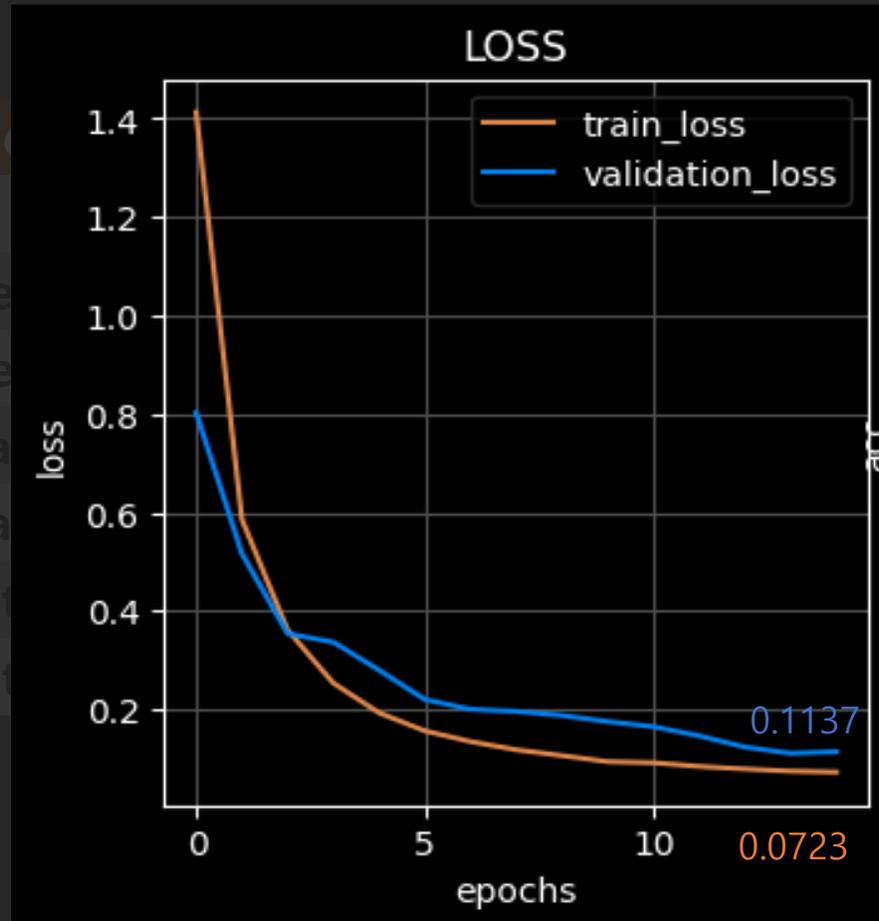
Basic





+ 코드 + 텍스트

1 Fine_Tuning



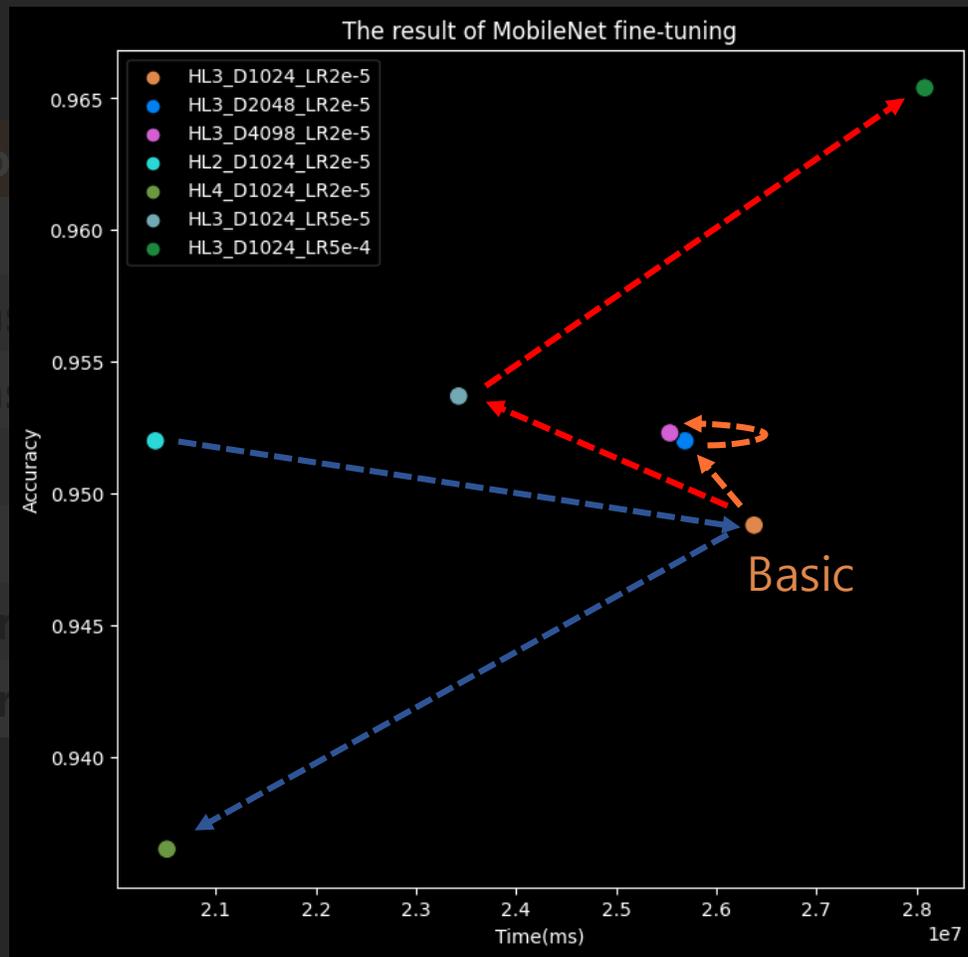
Model	g Time(ms)
Basic	76814
2048 dense	86338
4098 dense	35377
2 hidden la	95814
4 hidden la	11520
learning ra	25362
learning ra	82135

Final



+ 코드 + 텍스트

1 Fine_Tuning



Learning Rate	Accuracy	Learning Time(ms)
2e-5	95.20%	25686338
2e-5	95.23%	25535377
2e-5	93.65%	20511320
5e-5	95.37%	23425362
2e-4	96.54%	28082135

덴스 수 증가 = 정확도 증가

히든 레이어 수 증가 = 정확도 감소

러닝 레이트 증가 = 정확도 증가



+ 코드 + 텍스트

```
1 def food_table( "가곡초등학교" )
```

<https://www.foodsafetykorea.go.kr/main.do>

```
1 def crawling( Food_table_highschool )
```

```
C:\Users\admin\AppData\Local\Temp\ipykernel_23212\289921546.py:53: DeprecationWarning: use options instead of chrome_options
driver = webdriver.Chrome(service=Service(ChromeDriverManager().install()), chrome_options=chrome_options)
0%|          | 0/92 [00:00<?, ?it/s]C:\Users\admin\AppData\Local\Temp\ipykernel_23212\289921546.py:86: FutureWarning: elementwise comparison
if sch_nm in sch_list: continue
1%|█         | 1/92 [01:24<2:09:36, 84.80s/it]해운대여자고등학교 (부산광역시교육청) => 0:01:24.767540
2%|██        | 2/92 [02:07<1:30:15, 60.17s/it]행신고등학교 (경기도교육청) => 0:00:42.891749
3%|███       | 3/92 [02:50<1:17:38, 52.34s/it]향남고등학교 (경기도교육청) => 0:00:42.994973
4%|████      | 4/92 [03:16<1:01:26, 41.90s/it]향동고등학교 (경기도교육청) => 0:00:25.829281
5%|█████     | 5/92 [03:59<1:01:26, 42.38s/it]향일고등학교 (경기도교육청) => 0:00:43.172267
7%|██████    | 6/92 [04:42<1:01:02, 42.59s/it]현대고등학교 (울산광역시교육청) => 0:00:42.913879
8%|███████   | 7/92 [05:38<1:06:28, 46.92s/it]현대고등학교 (서울특별시교육청) => 0:00:55.784657
9%|████████  | 8/92 [06:22<1:04:05, 45.78s/it]현대공업고등학교 (울산광역시교육청) => 0:00:43.274963
10%|█████████| 9/92 [08:13<1:31:49, 66.38s/it]현대청운고등학교 (울산광역시교육청) => 0:01:51.623220
11%|█████████| 10/92 [08:29<1:09:24, 50.79s/it]현도정보고등학교 (충청북도교육청) => 0:00:15.830094
12%|█████████| 11/92 [09:12<1:05:28, 48.51s/it]현서고등학교 (경상북도교육청) => 0:00:43.269493
13%|█████████| 12/92 [09:55<1:02:09, 46.62s/it]현암고등학교 (경기도교육청) => 0:00:42.243616
14%|█████████| 13/92 [11:30<1:20:47, 61.36s/it]현일고등학교 (경상북도교육청) => 0:01:35.201380
15%|█████████| 14/92 [12:53<1:28:09, 67.81s/it]현천고등학교 (강원도교육청) => 0:01:22.704597
16%|█████████| 15/92 [14:20<1:34:33, 73.69s/it]현풍고등학교 (대구광역시교육청) => 0:01:27.262046
17%|█████████| 16/92 [15:25<1:30:05, 71.13s/it]현화고등학교 (경기도교육청) => 0:01:05.108297
18%|█████████| 17/92 [16:08<1:18:23, 62.71s/it]형성고등학교 (대구광역시교육청) => 0:00:43.100947
20%|█████████| 18/92 [17:24<1:22:02, 66.52s/it]형곡고등학교 (경상북도교육청) => 0:01:15.312032
21%|█████████| 19/92 [18:59<1:31:34, 75.26s/it]형석고등학교 (충청북도교육청) => 0:01:35.584475
22%|█████████| 20/92 [20:11<1:28:54, 74.09s/it]혜광고등학교 (부산광역시교육청) => 0:01:11.317203
23%|█████████| 21/92 [20:55<1:17:14, 65.28s/it]혜성여자고등학교 (서울특별시교육청) => 0:00:44.707305
24%|█████████| 22/92 [21:54<1:13:55, 63.37s/it]혜원여자고등학교 (서울특별시교육청) => 0:00:58.839262
25%|█████████| 23/92 [22:43<1:07:46, 58.93s/it]혜화여자고등학교 (서울특별시교육청) => 0:00:48.555860
26%|█████████| 24/92 [24:03<1:13:49, 65.14s/it]혜화여자고등학교 (부산광역시교육청) => 0:01:19.574095
27%|█████████| 25/92 [24:45<1:05:11, 58.38s/it]호계고등학교 (울산광역시교육청) => 0:00:42.561373
28%|█████████| 26/92 [26:31<1:19:47, 72.53s/it]호남고등학교 (전라북도교육청) => 0:01:45.519793
28%|█████████| 26/92 [26:33<1:07:24, 61.29s/it]호남삼육고등학교 (광주광역시교육청) error
```



+ 코드 + 텍스트

1 len(menu)

249238

1 split_words

["와", "과", "랑"]

[] 1 exception_menus

[["사과", "견과", "견과", "빙과", "모과", "약과", "제과", "유과", "용과", "무화과", "백향과", "하미과", "한과", "정과"], ["구프랑", "몽블랑"]]

1 menu_csv.to_csv ("test_csv.csv", encoding="euc-kr", index=False)

775913	울곡고등학교 (경상북도교육청)	2019.12.19	중식	기장밥. 세발나물(자율) 물미역/초고추장(자율)5.6.13. 짜꾸미떡볶음5.6.12...	기장밥 깍두기 설렁탕
775914	울곡고등학교 (경상북도교육청)	2019.12.23	석식	쌀밥 순대국밥(냉면기)5.6.8.9.10.13.16.18. 모듬야채겉절이13. 깍두...	순대국밥 배추겉절이 햄버거 깍두기
775915	울곡고등학교 (경상북도교육청)	2019.12.23	중식	찰흑미밥. 콩나물김치국5.6.9.13.18. 비빔짬면4.5.6.13. 쇠불고기5.6...	짬면 흑미밥 김치국 배추김치 쇠불고기
775916	울곡고등학교 (경상북도교육청)	2019.12.30	석식	자장면1.2.5.6.10.13.16.18. 후라이드 닭꼬치1.5.6.13.15. 단...	자장면 닭꼬치 김치볶음 단무지
775917	울곡고등학교 (경상북도교육청)	2019.12.30	중식	기장밥. 쇠고기미역국.5.6.13.16.18. 오향장육5.6.10.13.18. 부추...	기장밥 배추겉절이 쇠고기미역국 골뱅이무침 배추김치



+ 코드 + 텍스트

1 식품영양성분DB 가져오기

<https://various.foodsafetykorea.go.kr/nutrient/>

```
1 db_menus = []
2 with open(os.path.join(data_path, "db_menus.txt"), 'r') as f:
3     for line in f:
4         db_menus.append(line.replace("\n", ""))
5 len(db_menus)
```

787

```
1 #식품영양성분DB 메뉴 중 학교급식정보 메뉴에서 검출된 목록
2 with open(os.path.join(data_path, "include_db_menus.txt"), "w", encoding="euc-kr") as f:
3     for key, value in menus.items():
4         f.write(f"{key}: {value}\n")
```

714

+ 코드 + 텍스트

```
1 df [ df [ 'menu' ].str.contains ( r".*김치.*", regex=True) ]
```

[195601 rows x 4 columns]

df = 2660개중 401~899번째 학교의 식단표
[245179 rows x 4 columns]

```
1 df [ df [ 'menu' ].str.contains ( r".*단무지.*", regex=True) ]
```

[15797 rows x 4 columns]

```
1 df [ df [ 'menu' ].str.contains ( r".*오이.*무침.*", regex=True) ]
```

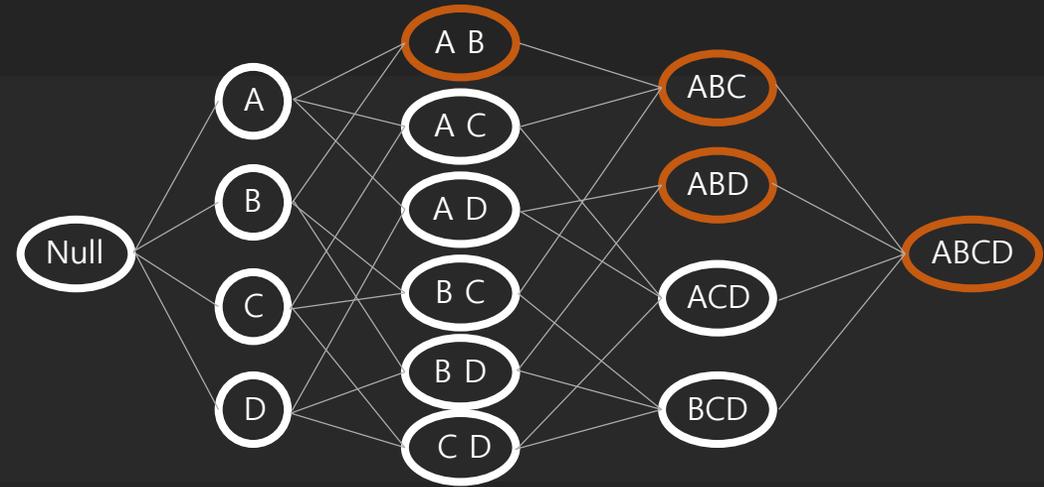
[13187 rows x 4 columns]

```
1 df [ df [ 'Menu' ] == "등심돈까스" ]
```

	sch_nm	menu_date	dt	menu
6678	미림여자정보과학고등학교 (서울특별시교육청)	2022.07.05	석식	찰쌀밥 미니 냉모밀 마파두부 생등심돈까스 단무지 요거타임
32195	목포마리아회고등학교 (전라남도교육청)	2022.06.29	중식	참치김치볶음밥 냉모밀 등심돈까스 소스 깍두기 구슬아이스크림 하트단무지
38483	목동고등학교 (서울특별시교육청)	2022.05.18	중식	후리가케밥 미니냉모밀 등심돈까스 통감자구이 깍두기 모찌롤케이크 양배추샐러드
49049	매곡고등학교 (충청북도교육청)	2020.08.11	석식	꼬마밥 미니냉모밀 양배추샐러드 치즈등심돈까스 소스 배추김치 수박
60098	마산내서여자고등학교 (경상남도교육청)	2021.10.01	석식	냉모밀 국수 아몬드스틱파이 등심돈까스 꼬들단무지 방울토마토12.

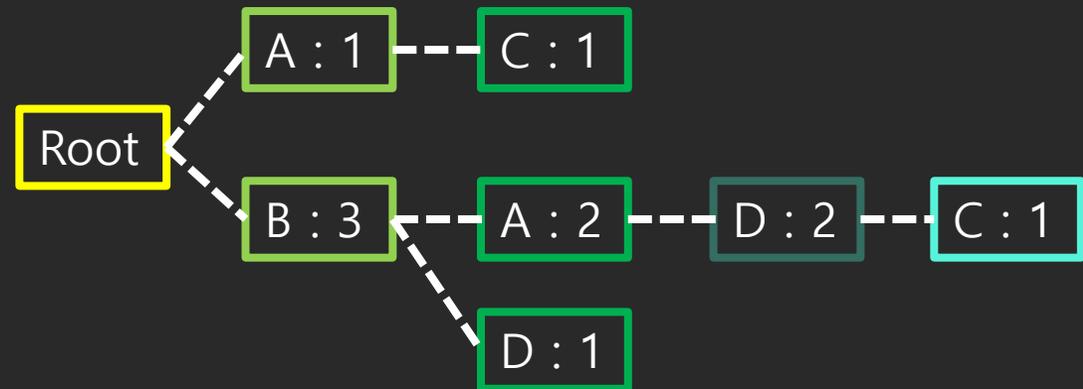
1 Apriori

- 빈발 아이템 집합
- 모든 가능한 연관 규칙
- 큰 데이터셋 => **성능저하**



1 FP-Growth

- FP - 트리 생성
- Apriori보다 빠른 속도
- 집합을 압축하여 메모리 사용량 감소
- 알고리즘 이해 및 구현 => **복잡**





+ 코드 + 텍스트

```
1 df [ df ['menu'].str.contains ( r".*김치.*", regex= True ) ]
```

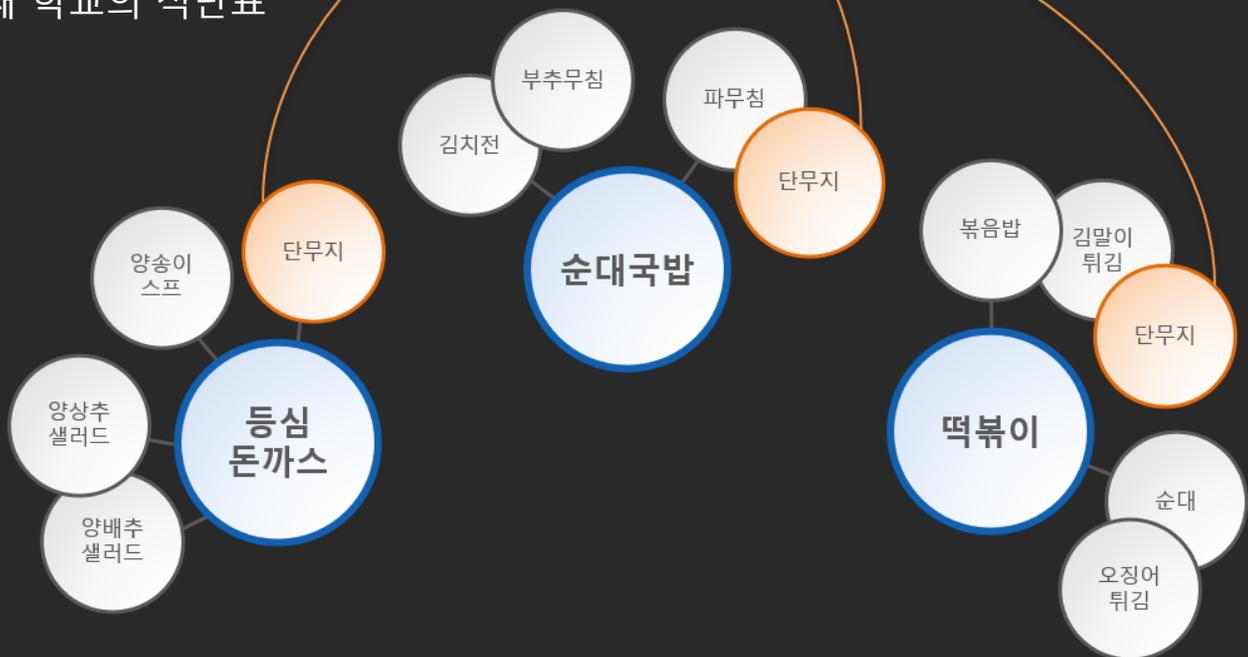
[195601 rows x 4 columns]

```
1 df [ df ['menu'].str.contains ( r".*단무지.*", regex= True ) ]
```

[15797 rows x 4 columns]

df = 2660개중 401~899번째 학교의 식단표
[245179 rows x 4 columns]

빈도 수 영향





+ 코드 + 텍스트

1 Confidence(신뢰도)

- 규칙의 강도와 신뢰성을 평가
- 값이 높을수록 **관련성**이 높음
- 아이템 간의 강한 **관련성**을 가진 규칙을 선택

(조건 아이템과 결과 아이템의 동시발생빈도)

(조건 아이템이 발생하는 빈도)

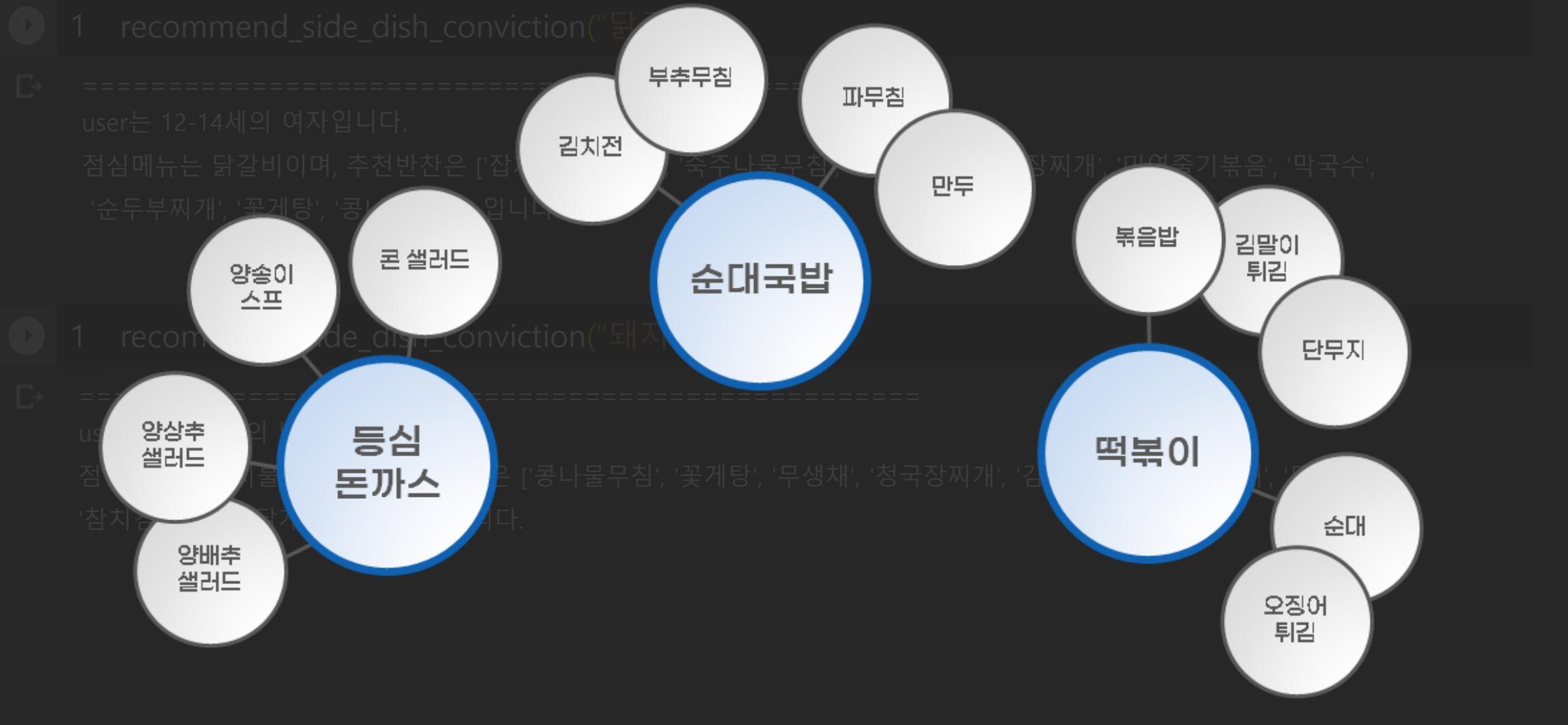
1 Conviction(확신도)

- 예측력과 독립성을 평가
- 값이 1에 가까울 수록 **독립성**이 높음
- 아이템 간의 **독립성**을 가진 규칙을 선택

(1 - Support(결과 아이템의 지지도))

(1 - Confidence(규칙의 신뢰도))

+ 코드 + 텍스트





+ 코드 + 텍스트

```
1 dinner=list ( recommend_dinner_conviction (lunch_menu) )
```

입력한 메뉴(순대국밥)를 통해 추천된 저녁 메뉴는 다음과 같습니다.

['쇠고기볶음', '굴비구이', '알밥', '달래된장찌개', '참치볶음밥', '참치김밥', '새우구이', '순두부김치찌개', '갈치조림', '등심돈가스']

```
1 저녁 군집분석
```

저녁군집분석결과 : 26 개, ['간장양념닭다리구이', '치즈돈가스', '오리탕', '부대찌개', '쇠갈비구이', '삼겹살고추장구이', '회냉면', '고등어석쇠구이', '치킨데리야끼', '짬면', '떡만두국', '콩국수', '콩치구이', '막국수', '간짜장', '잡채밥', '닭날개튀김', '고등어튀김', '간장양념치킨', '간장양념닭다리튀김', '후라이드치킨', '치킨까스', '양념치킨', '안심돈가스', '생선까스', '등심돈가스']

```
1 저녁추천 = 연관분석 ∩ 군집분석
```

교집합 : ['등심돈가스']

1 1인 1일 영양소 섭취 기준

성별	연령별	영양소	평균필요량	권장섭취량	
4	남자	19-29	에너지(kcal)	2600.0	0.0
13	여자	19-29	에너지(kcal)	2000.0	0.0
22	남자	19-29	탄수화물(g)	100.0	130.0
31	여자	19-29	탄수화물(g)	100.0	130.0
40	남자	19-29	지방(g)	0.0	0.0
49	여자	19-29	지방(g)	0.0	0.0
58	남자	19-29	단백질(g)	50.0	65.0
67	여자	19-29	단백질(g)	45.0	55.0

3대 영양소(탄수화물, 단백질, 지방)
& 에너지 섭취량(kcal) 고려

성별	연령별	영양소	하한(%)	상한(%)	
40	남자	19-29	지방(g)	15	30
49	여자	19-29	지방(g)	15	30

지방 섭취 상한량 추정

- 성별, 연령별 에너지 섭취량 중 지방의 상한비율(약 30%)
- 지방 1g 당 약 9kcal

⇒ 에너지섭취권장량 / 30

성별	연령별	에너지(kcal)	영양소	권장섭취량	
21	남자	19-29	2600.0	지방(g)	86.666667
66	여자	19-29	2000.0	지방(g)	66.666667

- 국민 1일 나트륨 섭취량, WHO기준 1.5배
- 최근 10년간, 19세 이상 성인 비만 유병률 증가



'총 당류, 나트륨' 추가

총 당류 상한량

- 적정섭취량 = 성별, 연령별 에너지 섭취권장량의 10%
- 당류 1g 당 약 4kcal

⇒ 에너지섭취권장량 / 40

나트륨 상한량

⇒ 모두 2g으로 통일

+ 코드 + 텍스트

```

1 food_nutrition .sort_values( '연도', ascending=False)
2 food_nutrition .drop_duplicates( '식품명' )
3 food_nutrition [ food_nutrition[ "지역 / 제조사" ] == "전국(대표)" ]
    
```

```

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 7704 entries, 0 to 7703
Data columns (total 100 columns):
#   Column                Non-Null Count  Dtype
---  ---                ---
0   NO                     7704 non-null   int64
1   SAMPLE_ID             7704 non-null   object
2   식품코드              7704 non-null   object
3   DB군                  7704 non-null   object
4   상용제품              7704 non-null   object
5   식품명                7704 non-null   object
6   연도                  7704 non-null   int64
7   지역 / 제조사         7704 non-null   object
8   채취시기              7704 non-null   object
9   식품대분류            7704 non-null   object
10  식품상세분류          7704 non-null   object
11  1회제공량             7704 non-null   int64
12  내용량_단위           7704 non-null   object
13  총내용량(g)           7704 non-null   object
14  총내용량(mL)          7704 non-null   object
15  에너지(kcal)          7704 non-null   object
16  수분(g)               7704 non-null   object
17  단백질(g)            7704 non-null   object
18  지방(g)               7704 non-null   object
19  탄수화물(g)           7704 non-null   object
20  총당류(g)             7704 non-null   object
96  회분(g)               7704 non-null   object
97  카페인(mg)           7704 non-null   object
98  성분표출처           7704 non-null   object
99  발행기관              7704 non-null   object
dtypes: int64(3), object(97)
memory usage: 5.9+ MB
    
```



```

<class 'pandas.core.frame.DataFrame'>
Int64Index: 793 entries, 7063 to 152
Data columns (total 100 columns):
#   Column                Non-Null Count  Dtype
---  ---                ---
0   NO                     793 non-null   int64
1   SAMPLE_ID             793 non-null   object
2   식품코드              793 non-null   object
3   DB군                  793 non-null   object
4   상용제품              793 non-null   object
5   식품명                793 non-null   object
6   연도                  793 non-null   int64
7   지역 / 제조사         793 non-null   object
8   채취시기              793 non-null   object
9   식품대분류            793 non-null   object
10  식품상세분류          793 non-null   object
11  1회제공량             793 non-null   int64
12  내용량_단위           793 non-null   object
13  총내용량(g)           793 non-null   object
14  총내용량(mL)          793 non-null   object
15  에너지(kcal)          793 non-null   object
16  수분(g)               793 non-null   object
17  단백질(g)            793 non-null   object
18  지방(g)               793 non-null   object
19  탄수화물(g)           793 non-null   object
20  총당류(g)             793 non-null   object
96  회분(g)               793 non-null   object
97  카페인(mg)           793 non-null   object
98  성분표출처           793 non-null   object
99  발행기관              793 non-null   object
dtypes: int64(3), object(97)
memory usage: 625.7+ KB
    
```

식품 영양 정보

- 연도별로 내림차순 정렬
- 식품명 중복 제거
- 지역/제조사를 전국(대표)로 한정

전처리 과정

- 7704개의 메뉴 영양 정보 중 필요한 793개의 영양 정보 선정



+ 코드 + 텍스트

1 food_nutrition [food_nutrition ["식품명", "에너지(kcal)", "탄수화물(g)", "단백질(g)", "지방(g)", "총당류(g)", "나트륨(g)"]]

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 793 entries, 7063 to 152
Data columns (total 100 columns):
#   Column                Non-Null Count  Dtype
---  ---                -
0   NO                    793 non-null    int64
1   SAMPLE_ID            793 non-null    object
2   식품코드             793 non-null    object
3   DB군                 793 non-null    object
4   상용제품             793 non-null    object
5   식품명              793 non-null    object
6   연도                 793 non-null    int64
7   지역 / 제조사        793 non-null    object
8   채취시기             793 non-null    object
9   식품대문류           793 non-null    object
10  식품상세문류         793 non-null    object
11  1회제공량            793 non-null    int64
12  내용량_단위           793 non-null    object
13  총내용량(g)          793 non-null    object
14  총내용량(mL)         793 non-null    object
15  에너지(kcal)         793 non-null    object
16  수분(g)              793 non-null    object
17  단백질(g)           793 non-null    object
18  지방(g)              793 non-null    object
19  탄수화물(g)          793 non-null    object
20  총당류(g)            793 non-null    object
...
96  회분(g)              793 non-null    object
97  카페인(mg)           793 non-null    object
98  성분표출처           793 non-null    object
99  발행기관             793 non-null    object
dtypes: int64(3), object(97)
memory usage: 625.7+ KB
```



```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 793 entries, 7063 to 152
Data columns (total 8 columns):
#   Column                Non-Null Count  Dtype
---  ---                -
0   식품명                793 non-null    object
1   에너지(kcal)          793 non-null    float64
2   탄수화물(g)           793 non-null    float64
3   단백질(g)             793 non-null    float64
4   지방(g)               793 non-null    float64
5   총당류(g)             793 non-null    float64
6   나트륨(g)             793 non-null    float64
7   클러스터명            0 non-null      object
dtypes: float64(6), object(2)
memory usage: 55.8+ KB
```



```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 218 entries, 7073 to 275
Data columns (total 8 columns):
#   Column                Non-Null Count  Dtype
---  ---                -
0   식품명                218 non-null    object
1   에너지(kcal)          218 non-null    float64
2   탄수화물(g)           218 non-null    float64
3   단백질(g)             218 non-null    float64
4   지방(g)               218 non-null    float64
5   총당류(g)             218 non-null    float64
6   나트륨(g)             218 non-null    float64
7   클러스터명            0 non-null      object
dtypes: float64(6), object(2)
memory usage: 15.3+ KB
```

차원 축소

- 식품명, 1회 제공량, 영양소 등
필요한 컬럼 추출

저녁 메인 메뉴 후보 선정

- 793개의 메뉴 중

임의로 218개의 추천용 메뉴 선정

+ 코드 + 텍스트

1 **def** remain_nutri("여자", "13"):

점심영양소 =====

- 에너지(kcal) : [673.09]
- 탄수화물(g) : [93.46]
- 단백질(g) : [28.56]
- 지방(g) : [20.56]
- 총당류(g) : [1.51]
- 나트륨(g) : [1.14]

반찬영양소 =====

- 에너지(kcal) : [167.83]
- 탄수화물(g) : [21.62]
- 단백질(g) : [6.18]
- 지방(g) : [6.29]
- 총당류(g) : [0.62]
- 나트륨(g) : [0.35]

전체영양소(점심+반찬) =====

- 에너지(kcal) : 840.92
- 탄수화물(g) : 115.08
- 단백질(g) : 34.74
- 지방(g) : 26.85
- 총당류(g) : 2.13
- 나트륨(g) : 1.48

기준영양소 =====

- 에너지(kcal) : 2000.0
- 탄수화물(g) : 130.0
- 단백질(g) : 55.0
- 지방(g) : 66.67
- 총당류(g) : 50.0
- 나트륨(g) : 2.0

잔여영양소 =====

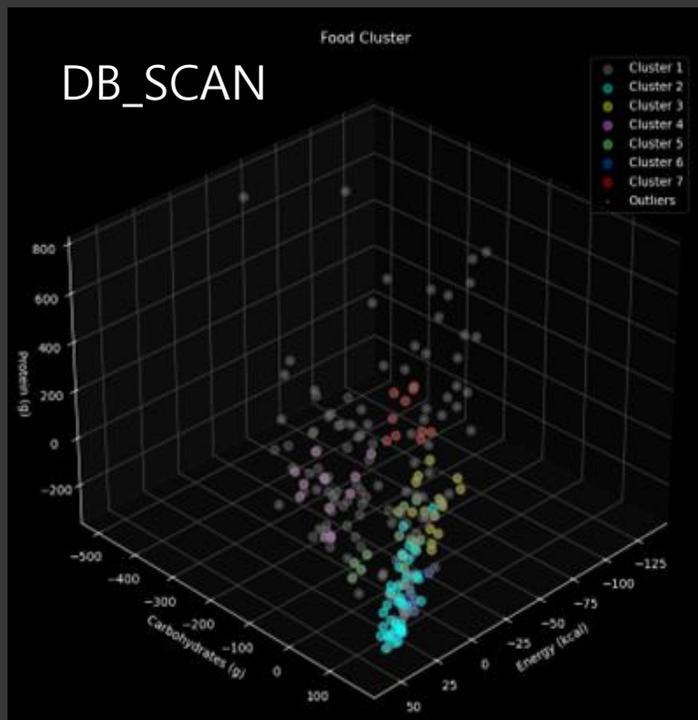
- 에너지(kcal) : 1159.08
- 탄수화물(g) : 14.92
- 단백질(g) : 20.26
- 지방(g) : 39.82
- 총당류(g) : 47.87
- 나트륨(g) : 0.52

$$\text{잔여영양소} = \text{기준영양소} - \text{전체영양소}$$

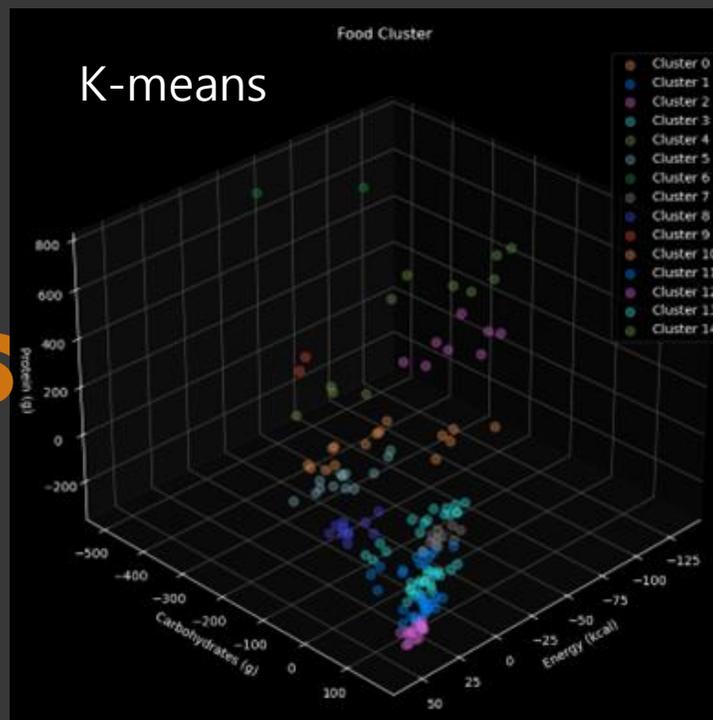
$$\text{전체영양소} = \text{점심영양소} + \text{반찬영양소}$$

기준 영양소 = 1인 1일 영양소 섭취 기준

1 DB_SCAN VS K-means



VS



K-means의 $n_clusters = 15$

DB_SCAN 선정 이유

① 연관분석과의 교집합 개수 고려

- 군집 내 **최소** 아이템 개수 설정
- 반면 K-means는 군집 개수 기반으로 군집 형성 ($n_clusters$)

② 영양소 기반 추천 정확도 고려

- K-means는 영양소가 비슷하지 않아도 같은 군집으로 형성
- DB_SCAN은 **밀도** 기반 군집으로 비슷하지 않은 데이터는 outlier



+ 코드 + 텍스트

1 DB_SCAN

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Epsilon	60	70	80	90	100	110	120	130	150	200	75	100	125	75	90	100	110	125	150	200
minPts	5	5	5	5	5	5	5	5	5	5	7	7	7	10	10	10	10	10	10	10
Clusters	9	9	10	10	10	10	10	9	9	9	9	9	9	5	7	7	7	8	8	9
Outliers	65	60	45	39	35	26	22	19	17	7	60	47	36	101	76	72	65	51	43	12

Epsilon이 높아질수록 Outlier는 감소

minPts가 높아질수록 묶이는 Cluster 수가 적어짐

> Cluster 수와 Outliers의 수, 각 Cluster의 군집요소의 밸런스가 중요

+ 코드 + 텍스트



1 DBSCAN(epsilon=110, min_samples=10)

```

Cluster 1 : 허미라이스 [401,93 76,84 14,47 4,08 0, 0,88409]
1 ['허미라이스', '표고버섯볶음밥', '콩나물국밥', '카레라이스', '치즈김밥', '채소볶음밥', '채소김밥', '참치볶음밥', '참치미요삭김밥', '참치김밥', '깡통밥', '짜장밥', '육회비빔밥', '유부초밥', '오징어덮밥', '오므라이스', '영양돌솥밥', '일밥', '계란볶음밥', '고추김밥', '순대국밥', '쇠고기주먹밥', '쇠고기볶음밥', '쇠고기김밥', '생선모듬초밥', '새우볶음밥', '비빔밥', '불고기덮밥', '볶음밥', '모듬회덮밥', '돌솥비빔밥', '돈까스김밥', '닭고기덮밥', '날치일밥', '낙지덮밥', '김치볶음밥', '김치김밥', '김밥']

Cluster 2 : 호박고추장찌개 [115,04 16,69 5,83 2,77 1,58 0,59073]
1 ['호박고추장찌개', '쌀 김치찌개', '해물매운탕', '콩비지찌개', '콩나물된장찌개', '수어탕', '황국장찌개', '참치김치찌개', '조개된장찌개', '오징어찌개', '오리백숙', '어묵김치찌개', '알탕', '순두부찌개', '순두부김치찌개', '쇠고기전골', '쇠고기된장찌개', '쇠고기 사브사브', '버섯전골', '버섯사브사브', '두부매운탕', '두부된장찌개', '두부고추장찌개', '두부고추장찌개', '된장찌개', '돼지고기찌개', '돼지고기김치찌개', '돌솥찌개', '대구 매운탕', '닭볶음탕', '알래된장찌개', '날이된장찌개', '낙지전골', '꽃게탕', '꿀치김치찌개', '김치전골', '갈치찌개']

Cluster 3 : 할박스테이크 [3,4737e+02 1,4380e+01 2,4200e+01 2,1450e+01 1,1000e-01 3,9528e-01]
1 ['할박스테이크', '닭발구이', '닭고기대파꼬치구이', '공비구이', '갈치구이', '가자미구이', '돼지고기산적', '키조개구이', '참치머리구이', '조기구이', '장어고추장구이', '오리불고기', '안성스테이크', '쇠불고기', '쇠곱창구이', '새우구이', '삼치구이', '뱅어포구이', '등심스테이크', '돼지불고기']

Cluster 4 : 깡통 [464,46 66,06 22,3 12,34 12,52 1,44606]
1 ['갈국수', '깡통', '짜장면', '경빈국수', '수제비', '비빔라면', '비빔국수', '볶음우동', '바지락갈국수', '물냉면', '라면', '떡라면', '달걀라면']

Cluster 5 : 죽순쇠고기볶음 [276,31 8,69 31,46 12,86 4,77 0,4925]
1 ['왕보쌈', '죽순쇠고기볶음', '장채', '유상술', '오징어볶음', '쇠고기채소볶음', '문어고추장볶음', '묵은지삼겹살볶음', '라볶이', '떡볶이', '돼지곱창순대볶음', '돼지고기김치볶음', '닭발볶음', '낙지볶음']

Cluster 6 : 닭계장 [142,98 9,88 19,17 2,97 4, 1,16818]
1 ['해물장국', '도가니탕', '닭계장', '꼬리곰탕', '곰탕', '새우탕', '육계장', '우거지곰탕', '얼갈탕', '선정국(선지해장국)']

Cluster 7 : 치킨짜스 [593, 44,02 27,91 33,92 4,15 0,77927]
1 ['닭날개튀김', '고등어튀김', '간장양념치킨', '간장양념다리튀김', '후라이드치킨', '치킨짜스', '영남치킨', '안심돈가스', '생선짜스', '등심돈가스']

```

DBSCAN 최종 모델

Epsilon = 110, minPts = 10

- 각 군집 내 아이템의 상호 유사성 이상적
- 65개의 Outliers

```

outliers ['제육덮밥', '잡채밥', '쇠머리국밥', '돼지머리국밥', '장어탕', '오리탕', '부대찌개', '내장탕', '감자탕', '꽂치구이', '고등어 석쇠구이', '고등어구이', '간장양념닭다리구이', '닭구이', '돼지갈비구이', '치킨데리야끼', '오리고기구이', '쇠갈비구이', '삼겹살구이', '삼겹살고추장구이', '회냉면', '콩국수', '쫄면', '잔치국수', '우동 (일식)', '미트볼 토마토 스파게티', '메밀국수', '막국수', '간짜장', '채소떡볶이', '참치 김치볶음', '제육볶음', '쇠고기볶음', '햄채소볶음', '돼지고기채소볶음', '닭볶음', '삼계탕', '닭곰탕', '순댓국', '갈비탕', '사골만두국', '떡만두국', '떡국', '등심돈가스', '닭강정', '깡풍기', '새우탕수육', '탕수육', '치즈돈가스', '달걀찜(우유)', '꼬막찜', '가오리찜', '달걀찜(새우젓)', '닭찜', '해물콩나물찜', '코다리찜', '족발', '아귀찜', '순대', '쇠고기수육', '쇠고기사태찜', '쇠갈비찜', '돼지고기 수육', '돼지갈비찜', '고등어조림', '건갈치조림', '갈치조림', '꽂치조림', '닭고추장조림', '닭조림', '코다리조림', '양미리조림', '삼치조림', '병어조림', '미트볼조림', '육회']

```



+ 코드 + 텍스트

▶ 1 잔여 영양소 and 클러스터 중앙값



	에너지(kcal)	탄수화물(g)	단백질(g)	지방(g)	총당류(g)	나트륨(g)
0	371.53	156.8	86.67	19.7	13.39	4.04

◀ 잔여 영양소 예시

	에너지(kcal)	탄수화물(g)	단백질(g)	지방(g)	총당류(g)	나트륨(g)
0	401.93	76.84	14.47	4.08	0.00	0.88409
1	115.04	16.69	5.83	2.77	1.58	0.59073
2	347.37	14.38	24.20	21.45	0.11	0.39528
3	464.46	66.06	22.30	12.34	12.52	1.44606
4	276.31	8.69	31.46	12.86	4.77	0.49250
...

◀ 각 클러스터의 중앙값이 클러스터를 대표

▶ 1 (각 클러스터 중앙값 - 잔여 영양소) .sum (axis=1)



	에너지(kcal)	탄수화물(g)	단백질(g)	지방(g)	총당류(g)	나트륨(g)	총합
0	30.40	79.96	72.20	15.62	13.39	3.15591	214.72591
1	256.49	140.11	80.84	16.93	11.81	3.44927	509.62927
2	24.16	142.42	62.47	1.75	13.28	3.64472	247.72472
3	92.93	90.74	64.37	7.36	0.87	2.59394	258.86394
4	95.22	148.11	55.21	6.84	8.62	3.54750	317.54750
...

총합 컬럼 추가

- 잔여 영양소의 총합
- Norm1

+ 코드 + 텍스트

menu_count : 1

n = 0

While True:

menu_count += n * Closest Cluster

n += 1

if menu_count >= 10:

break

	에너지(kcal)	탄수화물(g)	단백질(g)	지방(g)	총당류(g)	나트륨(g)	총합	Cluster	메뉴개수
9	40.12	137.38	59.44	3.61	2.39	3.56763	246.50763	미트볼조림	1
10	41.13	143.79	37.87	10.46	11.46	2.93494	247.64494	닭곰탕	1
11	24.16	142.42	62.47	1.75	13.28	3.64472	247.72472		3 20
12	92.93	90.74	64.37	7.36	0.87	2.59394	258.86394		4 13
13	42.76	145.80	50.37	4.19	12.99	3.13466	259.24466	쇠고기사태찜	1
14	78.56	111.76	62.99	0.23	5.44	2.73514	261.71514	닭강정	1
15	100.46	73.14	73.14	6.29	3.09	2.87541	266.83541	산도리	1
16	52.44	146.72	51.77	7.42	13.07	2.74521	274.16521	내장탕	1
17	51.52	155.00	49.57	1.43	13.39	3.36693	274.27693	닭찜	1
18	92.40	134.23	52.92	6.82	4.41	3.22520	294.00520	등심돈가스	1
19	147.76	142.69	80.29	15.94	1.45	3.50203	295.77203	채소떡볶이	1
20	137.93	67.09	72.57	9.23	11.81	2.05842	300.68842	순댓국	1
21	127.39	150.09	1.38	4.89	13.06	3.94819	301.35819	쇠고기수육	1
22	77.74	46.74	59.62	3.35	12.64	3.33499	302.62499	닭조림	1
23	95.22	148.11	55.21	6.84	8.62	3.54750	317.54750		5 14
24	115.52	135.95	64.92	10.19	11.15	3.22310	340.95310	쇠갈비찜	1
25	116.45	144.98	56.30	15.77	12.45	2.78807	348.73807	제육볶음	1
26	216.48	42.23	62.66	15.96	12.82	2.80595	352.95595	떡국	1
27	138.74	130.77	73.38	11.31	0.52	3.05546	357.77546	새우탕수육	1
28	132.32	129.76	70.84	12.17	13.39	2.64776	361.12776	쇠머리국밥	1

- ← 4th Closest : +1
- ← 3rd Closest : +1
- ← 2nd Closest : +1
- ← 1st Closest : +1
- ← 5th Closest : +14



+ 코드 + 텍스트

1 Print(아쉬운점, 향후과제)

- 영양 불균형 관련 통계 분석 부재
- 전이학습모델 및 Yolo와의 비교 미진행
- 데이터 신뢰성 및 검수 과정
- 식단 데이터 한정성
- 이미지 분류 카테고리 감소
- 분석 결과의 시각화 제약



+ 코드 + 텍스트

1 참고자료

김성아 등 4인, 『서울시민 만성질환 실태와 식생활 위험요인 분석』, 서울연구원(2022)

- <https://www.si.re.kr/node/65499>

식품안전나라 < 우리학교 급식식단 >

- https://www.foodsafetykorea.go.kr/portal/sensuousmenu/schoolMealsDetail.do?menu_grp=MENU_NEW03&menu_no

만개의 레시피

- <https://www.10000recipe.com/index.html>

식품의약품안전처 < 식품영양성분 데이터베이스 >

- <https://various.foodsafetykorea.go.kr/nutrient/>

보건복지부 < 2020 한국인 영양소 섭취기준 요약표 >

- https://www.mohw.go.kr/react/jb/sjb030301vw.jsp?PAR_MENU_ID=03&MENU_ID=032901&CONT_SEQ=370012

서울 열린데이터 광장, 서울시 지정·인증업소 현황

- <https://data.seoul.go.kr/dataList/OA-2741/S/1/datasetView.do>



+ 코드 + 텍스트

1 참고자료

Kakao I cloud <Vision>

- <https://www.kakaoicloud.com/service/detail/6-28>

AI 영양기반 플랫폼 <뉴트로지니>

- <https://nutrigeni.co.kr/>

두잉랩 <푸드렌즈>

- <https://www.foodlens.ai/>

케어워드 <필리>

- <https://pilly.kr/>

Velog <01. 과거의 추천시스템 : 연관상품 추천, Apriori 알고리즘, FP-Growth>

- <https://velog.io/@ttogle918/Recommendation-01>

Data Mining & Quality Analytics Lab <Sequential Pattern Mining Algorithms>

- <http://dmqm.korea.ac.kr/activity/seminar/237>