

주류 속성을 활용한 전통주 추천 알고리즘

CONTENTS

STEP1) 프로젝트 및 데이터 설명

STEP2) 대표 술 좌표 생성 알고리즘

STEP3) 가성비 필터링 알고리즘

STEP4) 도수 설정 알고리즘

STEP5) 중요요소 가중치 알고리즘

Step 1, 프로젝트 및 데이터 설명



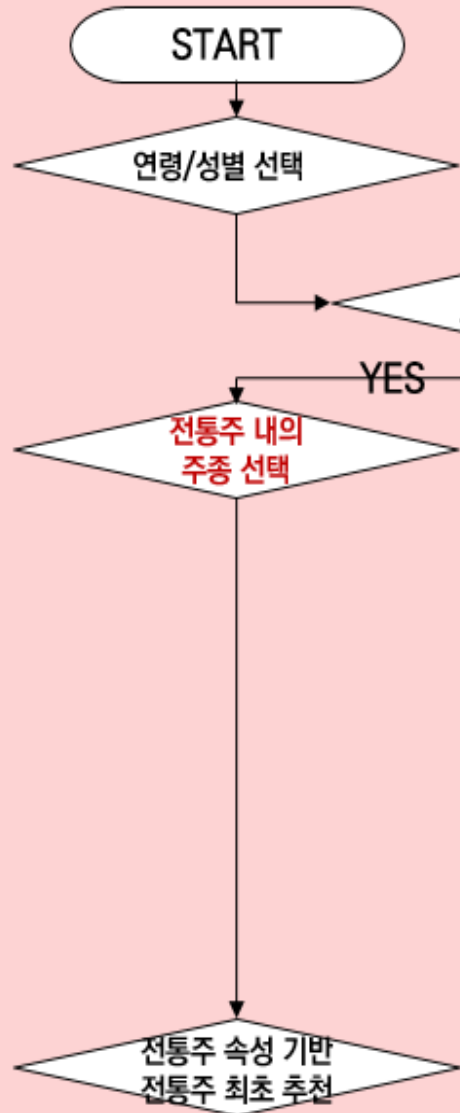
프로젝트 설명



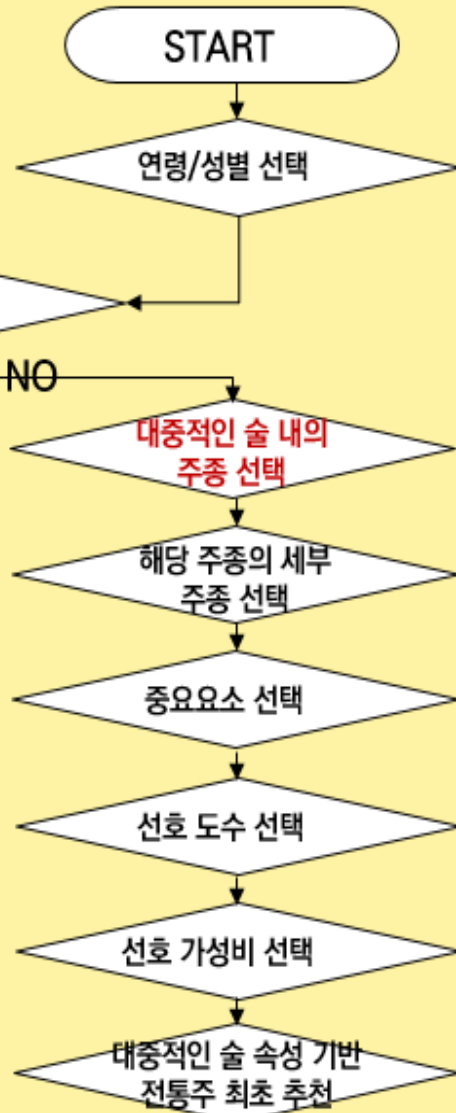
사용자가 전통주 추천 서비스에서 **주류 선호도 설문**을 하면
해당 설문 내용을 바탕으로 **4가지 알고리즘을 적용**하여 **개인 맞춤형 전통주 추천**

프로젝트 설명 - Service Flow

서비스1: 전통주 속성 기반 추천 서비스



서비스2: 대중적인 술 속성 기반 추천 서비스



서비스1. 선호하는 전통주가 있는 경우

- 해당 전통주와 주류 속성이 가장 유사한 전통주 추천

서비스2. 선호하는 전통주가 없는 경우

- 사용자가 선호하는 술, 주종, 중요요소, 도수, 가성비 기반 추천

데이터 설명



INPUT

사용자 선택할 전통주DB



OUTPUT

최종 추천할 전통주DB



USER

사용자의 설문 DB

Input 데이터 설명



INPUT

사용자 선택할 전통주DB

| 컬럼 | 컬럼 설명 | 세부 설명 |
|-------|------------------------|---|
| 상품명 | 상품의 이름 | - |
| 주종 | 상품의 주종 | 소주, 증류주, 맥주, 와인, 탁주, 약주, 과실주 |
| 세부 주종 | 상품의 세부 주종 (맥주와 와인만 존재) | 라거, 에일, 레드(스틸), 화이트(스틸), 화이트(스파클링), NAN |
| 절대 도수 | 상품의 도수 | - |
| 상대 도수 | 주종 내에서 해당 상품의 도수 정도 | 0 ~ 5 |
| 단맛 | 단맛의 정도 | 1 ~ 4 |
| 바디감 | 바디감의 정도 | 1 ~ 5 |
| 신맛 | 신맛의 정도 | 1 ~ 5, NAN |
| 청량감 | 청량감의 정도 | 1 ~ 5, NAN |
| 향의 강도 | 향의 강도의 정도 | 1 ~ 5, NAN |

사용자가 평소 좋아하는 술에 대한 데이터

Output 데이터 설명



OUTPUT

최종 추천할 전통주DB

| 컬럼 | 컬럼 설명 | 세부 설명 |
|---------|---------------------|------------------------------|
| 상품명 | 상품의 이름 | - |
| 주종 | 상품의 주종 | 소주, 증류주, 맥주, 와인, 탁주, 약주, 과실주 |
| 절대 도수 | 상품의 도수 | - |
| 상대 도수 | 주종 내에서 해당 상품의 도수 정도 | 0 ~ 5 |
| 용량 | 상품의 용량 | - |
| 가격 | 상품의 가격 | - |
| 가성비 | 상품의 가성비 | (용량 X 도수) / 가격 |
| 맥주 가성비 | 맥주정도의 가성비 여부 | 0(맥주보다 안 좋음), 1(맥주보다 좋음) |
| 검색용 상품명 | 검색 시 상품명 | 상품명으로 검색해서는 나오지 않는 상품 |
| 판매처 | 상품 판매처 | - |
| 단맛 | 단맛의 정도 | 1 ~ 5, NAN |
| 바디감 | 바디감의 정도 | 1 ~ 5, NAN |
| 신맛 | 신맛의 정도 | 1 ~ 5, NAN |
| 청량감 | 청량감의 정도 | 1 ~ 5, NAN |
| 향의 강도 | 향의 강도의 정도 | 1 ~ 5, NAN |
| 향의 종류 | 향의 종류 | 총 46개의 향의 종류 |
| 사진 | 상품의 사진 | - |

사용자로부터 입력 받은 정보를 바탕으로 최종 추천할 전통주 데이터

User 데이터 설명



USER

사용자의 설문 DB

| 컬럼 | 컬럼 설명 |
|----------------|------------------|
| unique_id | 고유 id |
| date_joined | 설문 참여 일자 |
| age | 나이 |
| sex | 성별 |
| alc_type | 선호하는 주종 |
| golm | 평소 즐겨 마시는 술 |
| factor | 술에 있어서 선호하는 요소 |
| scent | 선호하는 향 |
| alc_range_bool | 선호하는 도수 |
| CE_good_bool | 선호하는 가성비 (맥주 기준) |
| golajum | 최종 추천 술 |
| login_id | 로그인 id |
| login_pw | 로그인 pw |

사용자의 설문 내용이 담겨 있는 데이터

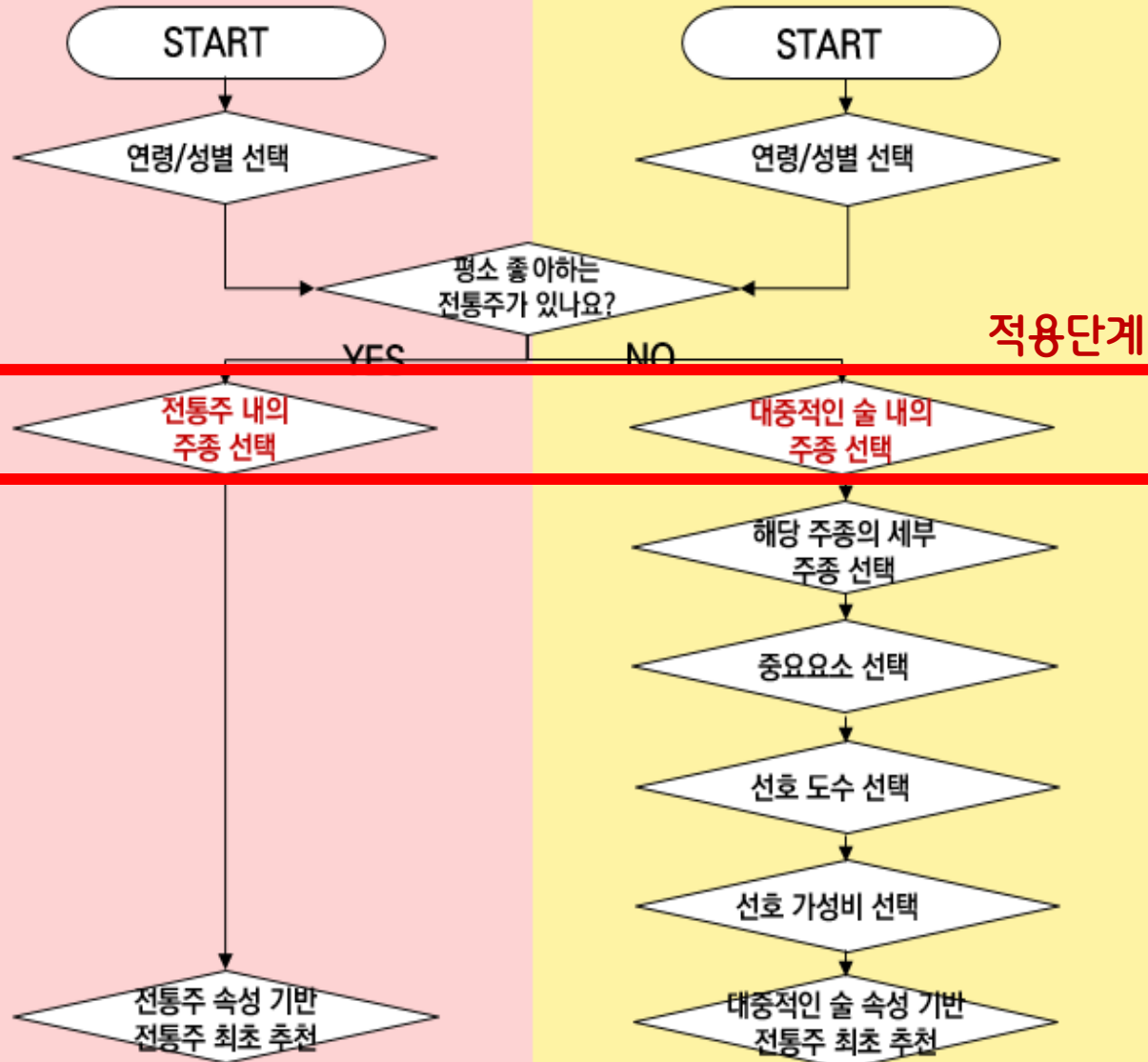
Step 2, 대표 술 좌표 생성 알고리즘



대표 술 좌표 생성 알고리즘 설명

서비스1: 전통주 속성 기반 추천 서비스

서비스2: 대중적인 술 속성 기반 추천 서비스



대표 술 좌표 생성 알고리즘

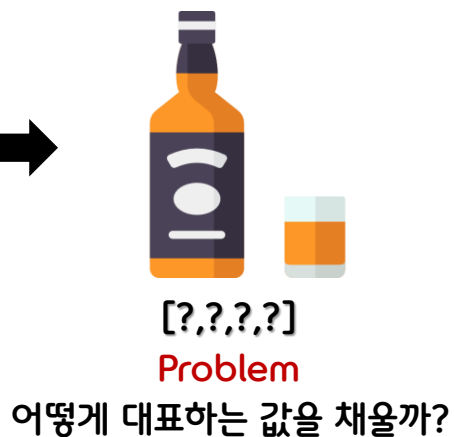
- 사용자가 선호하는 술을 선택하는 단계에 적용
- 선호하는 술 선택 시 해당 술들을 대표하는 값을 생성
- 해당 값을 기준으로 이후의 알고리즘을 적용

대표 술 좌표 생성 기준

사용자의 평소 선호술



가상의 술



Solution

실제 양조장의 실제 술 제작방식을 참고

- 양조장에서는 술을 만들 때, 원하는 특성의 평균값을 활용하여 최적의 조합을 찾는다
- 사용자가 선택한 3가지 술의 특성을 기반으로 가상의 술을 만들 때, 각 술의 특성값을 평균하여 최적의 특성을 도출

대표 술 좌표 생성 알고리즘

1. 대표 술 좌표 생성 알고리즘

사용자가 평소 선호하는 주류 상품들의 속성을 파악하여
사용자의 취향을 대표할 수 있는 가상의 상품을 생성

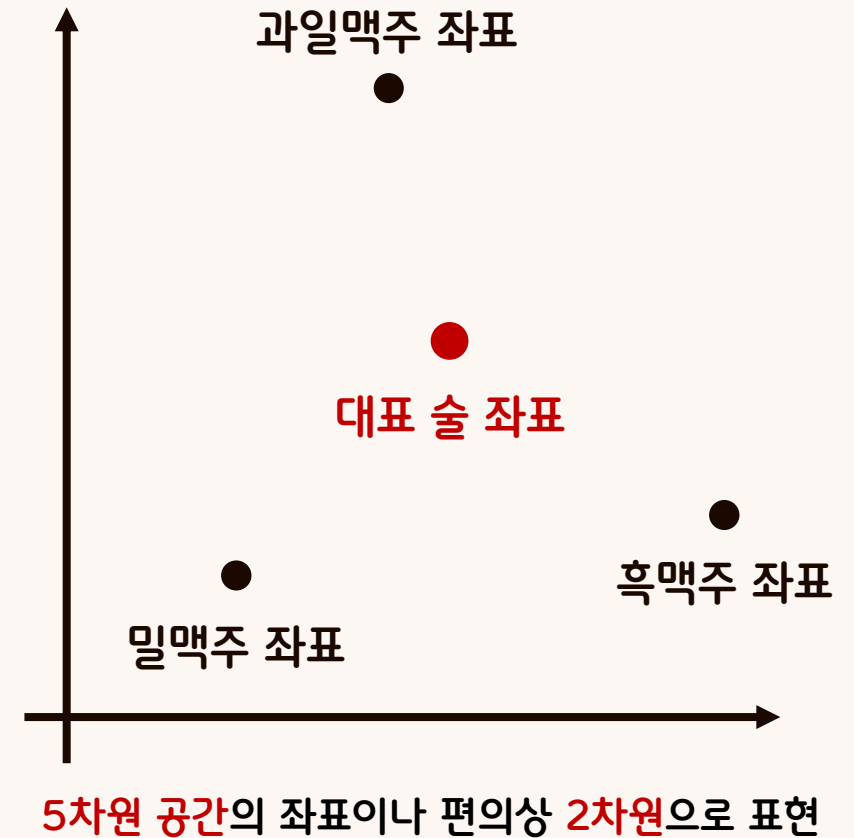
EX) 사용자의 입력

| 선호 주종 | 선호 상품명 | 대표 술 좌표 |
|-------|----------------|-------------------------|
| 맥주 | 과일맥주, 밀맥주, 흑맥주 | [3.3, 2.3, 3.3, 3, 1.3] |

적용 로직

각 속성의 평균값을 계산하여 대표값을 설정

알고리즘 실행 결과



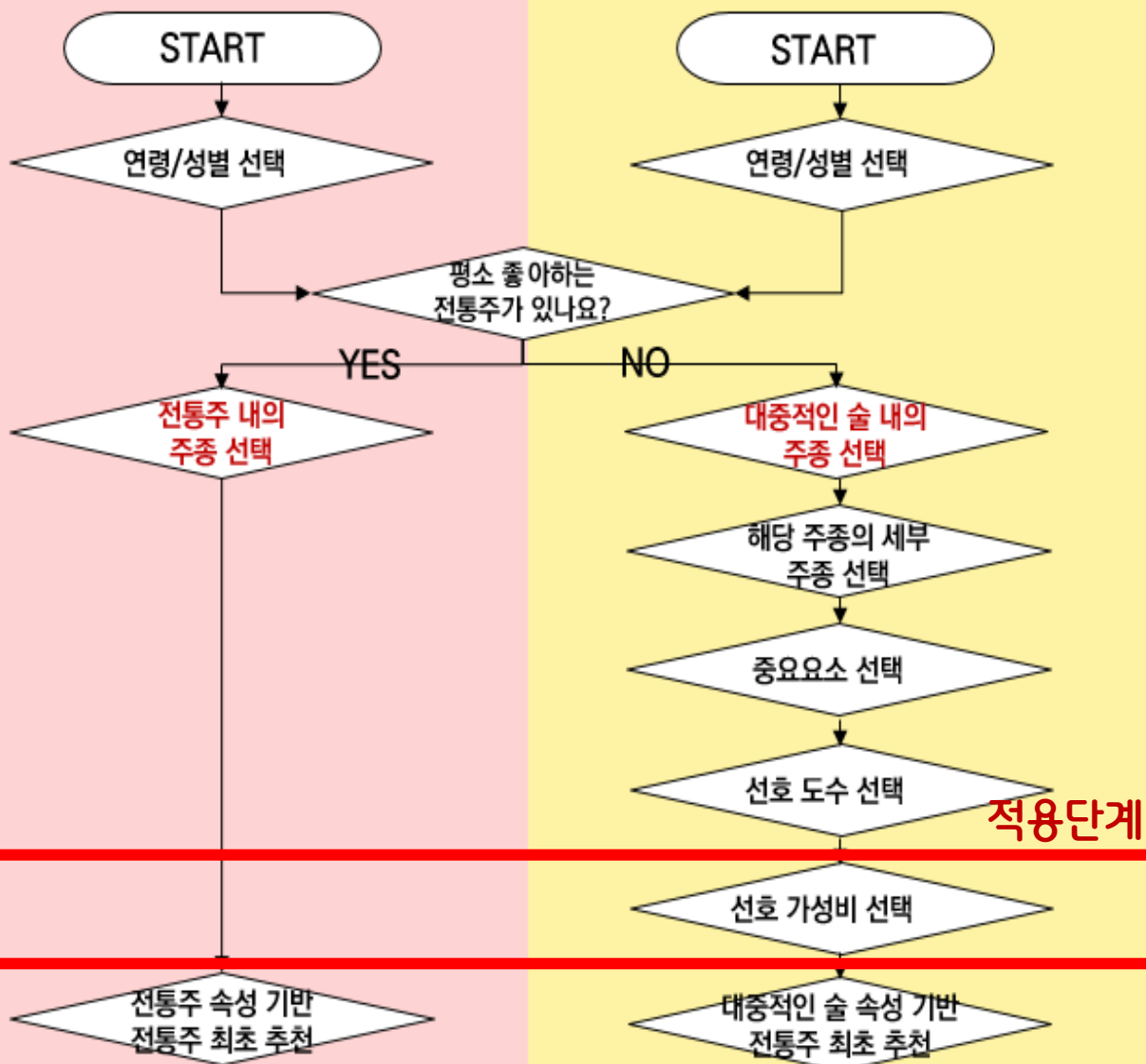
Step 3, 가성비 필터링 알고리즘



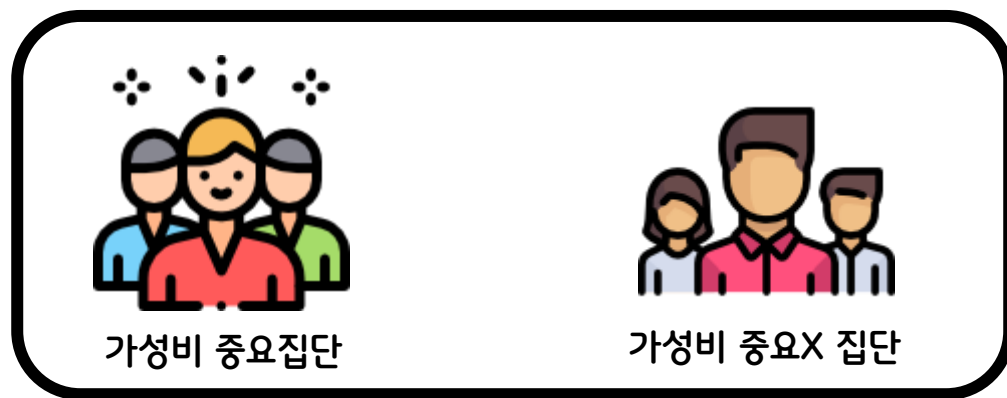
가성비 필터링 알고리즘 적용단계

서비스1: 전통주 속성 기반 추천 서비스

서비스2: 대중적인 술 속성 기반 추천 서비스



가성비 필터링 알고리즘



각 집단에 대해 다음과 같은 추천 방식을 적용

(1) 가성비가 중요하지 않은 집단

-> 가성비에 상관 없이 **모든 술을 추천**

(2) 가성비를 중요하게 생각하는 집단

-> 일정 수준의 가성비를 제시하고 **그 이상의 술 추천**

가성비 기준 선정

Problem



어떤 술이 가성비 기준?



Solution



1. 대중적인 술



2. 적절한 필터링 비율

사람들이 **가성비의 기준을 이해**하고 해당 질문에 답변하기 위해 **가성비의 기준이 되는 술**은 모두가 알고 있는 **대중적인 술**이어야 한다.

선택된 술을 기준으로 우리의 **Output 데이터**에서 **필터링**을 해야 하기에 **필터링의 정도가 30% ~ 70%** 정도 되어야 함

해당 조건을 바탕으로 **편의점 캔맥주**를 가성비 기준 선정

가성비 필터링 알고리즘

2. 가성비 필터링 알고리즘

사용자가 선호하는 가성비에 따라 추천 대상 전통주를 필터링

EX) 사용자의 입력

선호 가성비

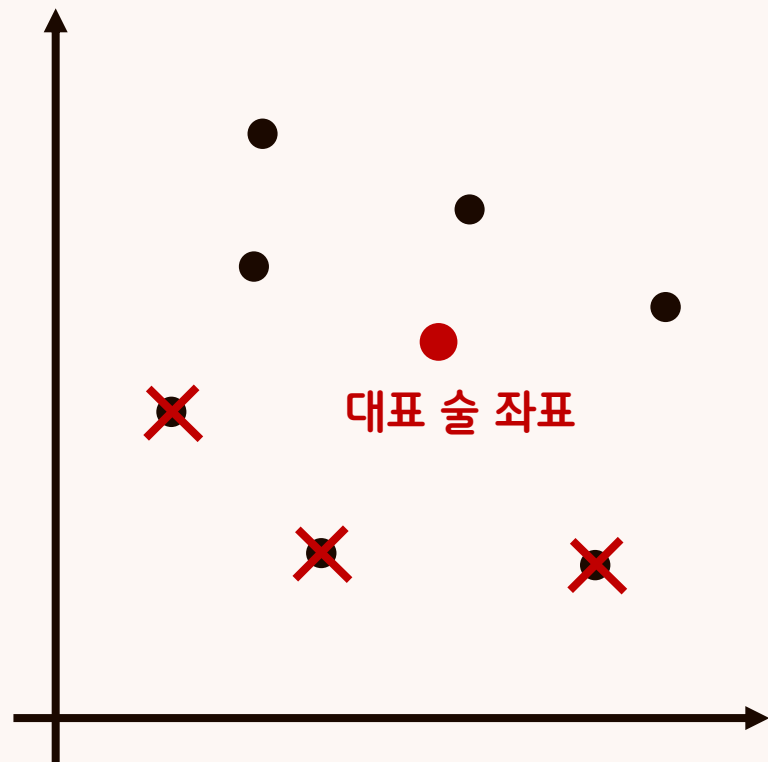
가성비 중요 ○

적용 로직

가성비의 좋고 나쁨의 기준은 편의점 맥주로 설정하여 질문
CASS 500ml, 2,500원, alc 4.5%

알고리즘 실행 결과

● Output 데이터



편의점 맥주보다 가성비가
떨어지는 제품 최종 추천 리스트에서 삭제

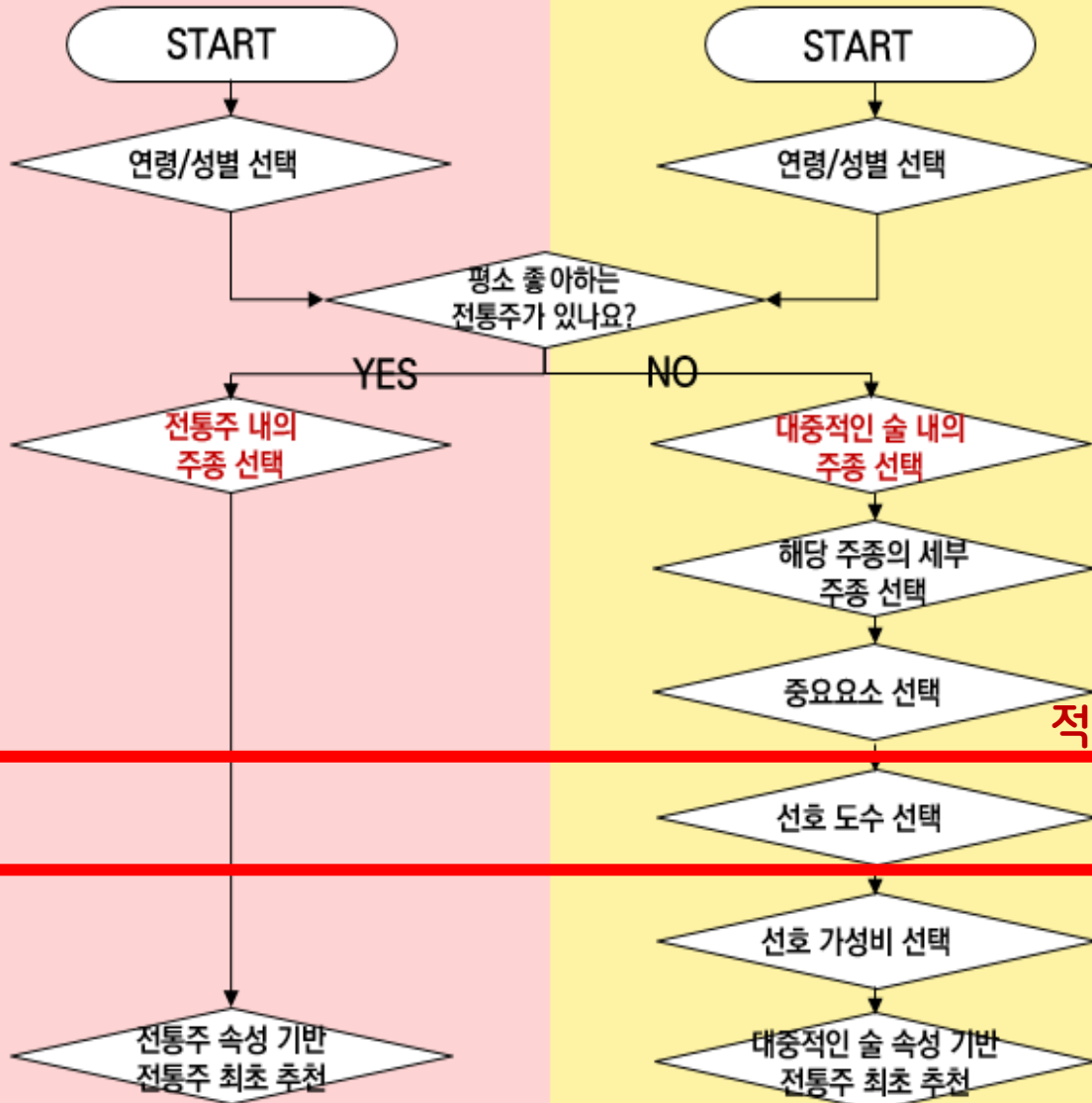
Step 4, 도수 설정 일고리즘



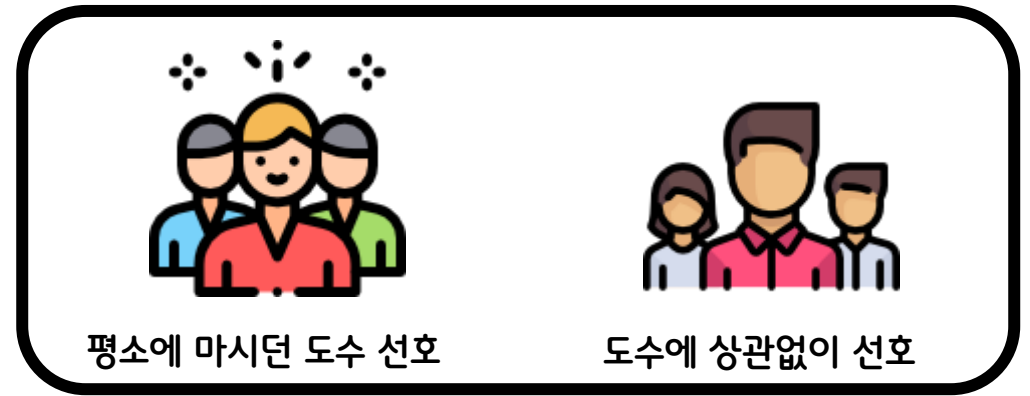
도수 설정 알고리즘 적용단계

서비스1: 전통주 속성 기반 추천 서비스

서비스2: 대중적인 술 속성 기반 추천 서비스



가성비 필터링 알고리즘

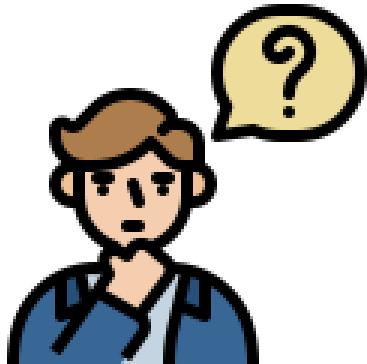


각 집단에 대해 다음과 같은 추천 방식을 적용

- (1) 평소에 마시던 도수를 선호하는 집단
-> 평소 마시던 도수에서 **3도 이내**의 술 추천
- (2) 도수에 상관없는 집단
-> **모든 술** 추천

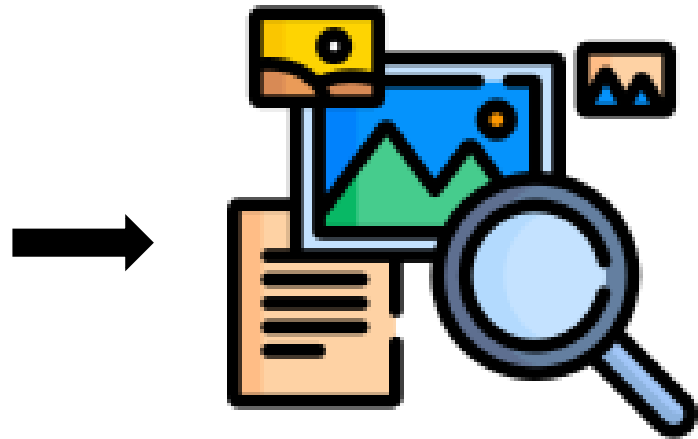
추천 도수 선정 기준

Problem



몇도까지 평소 마시던 도수 정의

Solution



선행 연구자료 참고

$\pm 3\%$

(3도 내의 범위 추천)

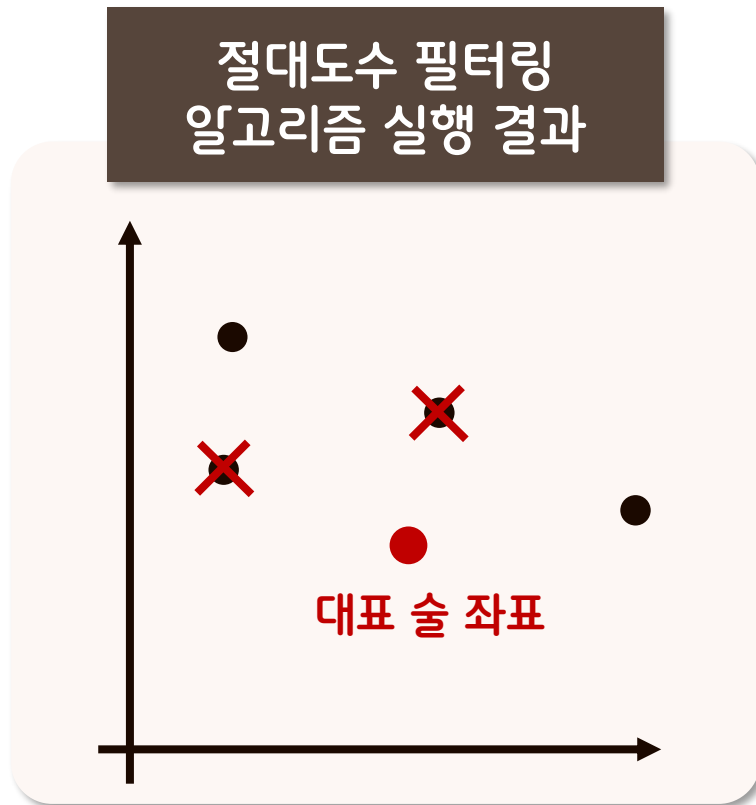
평소 마시던 술의 도수를 몇 도로 설정할 것인가에 대한 정의가 필요했고
이를 해결하기 위해 기존에 있던 술 추천 서비스를 참고하여 $\pm 3\%$ 로 설정하기로 정의

도수 설정 알고리즘

3. 도수 설정 알고리즘

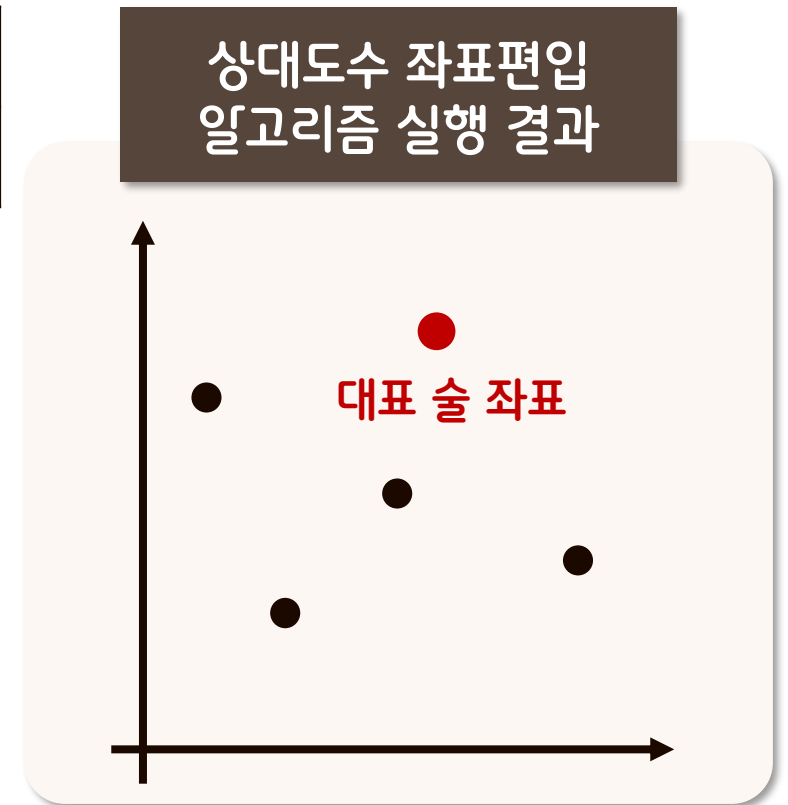
사용자가 선호하는 도수에 따라 '도수' 컬럼을 추천 알고리즘에 적용

| 선호 도수 |
|------------------|
| 평소 마시던 도수를 선호 |



선호하는 도수에서 $\pm 3\%$ 이내의 전통주 필터링

| 선호 도수 |
|----------|
| 모든 도수 선호 |



상대도수 개념을 추가하여 새롭게 좌표 재구성

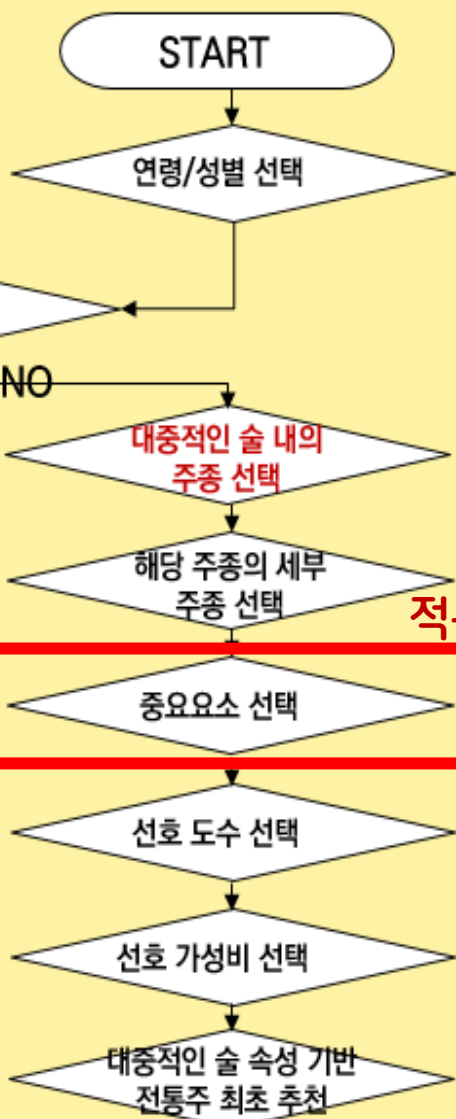
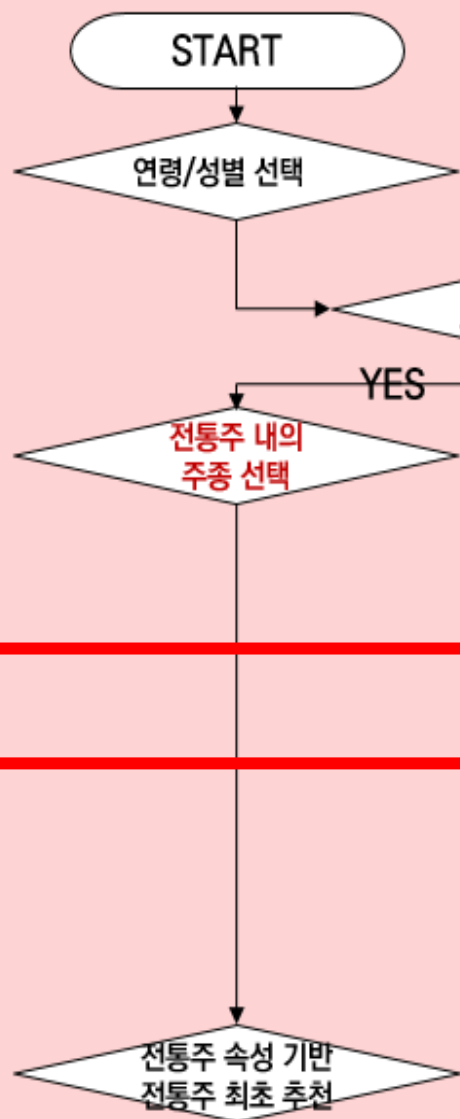
Step 5, 중요요소 가중치 일고리즘



중요요소 가중치 알고리즘 적용단계

서비스1: 전통주 속성 기반 추천 서비스

서비스2: 대중적인 술 속성 기반 추천 서비스



적용단계



중요요소

향의 종류 향의 강도 단맛 신맛 청량감 바디감

해당 요소 중 사용자가 중요하게 생각하는 요소에
가중치를 부여하여 더 잘 추천될 수 있도록 설정

중요요소 가중치 적용 공식



향의 종류 선호

$$\text{weight} = \max(0.2, 1 - \text{공통향의개수} / \text{전통주향의개수})$$



이외 요소 선호

향의 강도, 단맛, 신맛, 청량감, 바디감

$$\text{weight} = (6 - \text{중요요소 값}) / 5$$

가중치에 대한 명확한 공식이 없어 초기에는 임의로 설정
해당 가중치는 서비스 출시 이후 **사용자의 반응에 따라 조정할** 예정

중요요소 가중치 적용 공식

4. 중요요소 가중치 알고리즘

사용자가 중요하게 생각하는 요소의 값이 유사할수록
가중치를 부여해 추천 대상과의 유클리드 거리를 좁힘

EX) 사용자의 입력

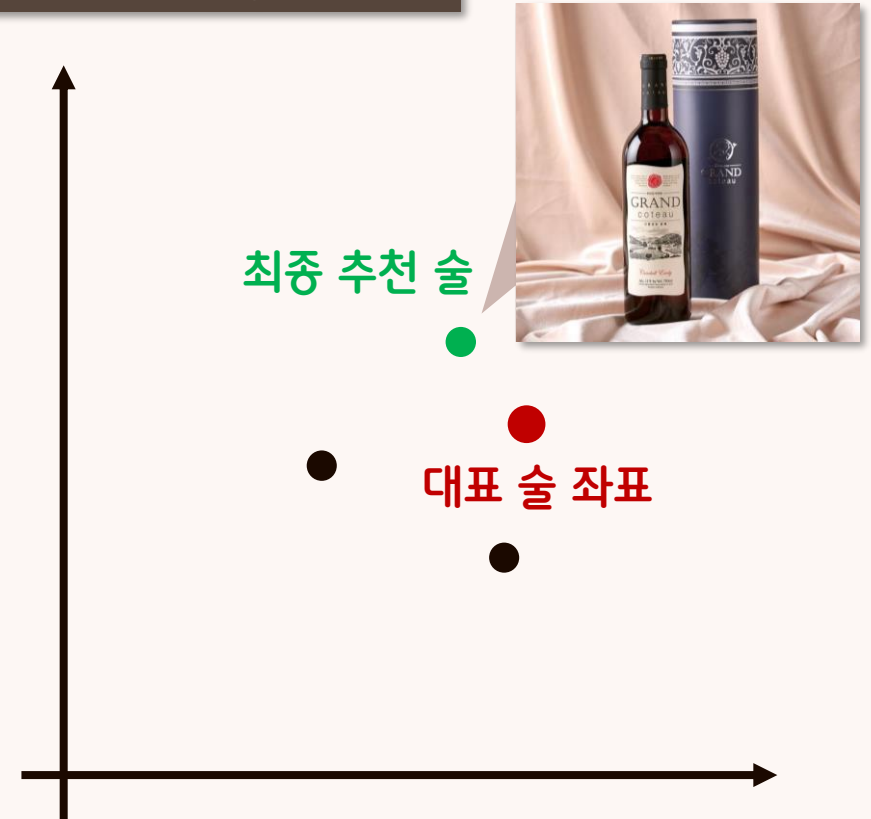
| 중요 요소 | 선호하는 향 |
|-------|-----------|
| 향의 종류 | [시트러스, 꽃] |

적용 로직

두 가지 경우로 분류해 가중치 계산

- ① 향의 종류를 선택한 사용자
- ② 그 외의 속성을 선택한 사용자

알고리즘 실행 결과



가중치를 적용하여 대표 술 좌표와 가장 가까운 전통주 추천