



대전 시내 택시 기사의 이동위치 추천



팀 캡틴즈 (이수민, 나영은, 이지현)

발표자 : 나영은

TABLE OF CONTENT

01 문제 배경 및 목적

02 활용 데이터 및
전처리 과정

03 EDA 시각화 진행

04 클러스터링 및
모델 알고리즘 소개

05 시뮬레이션

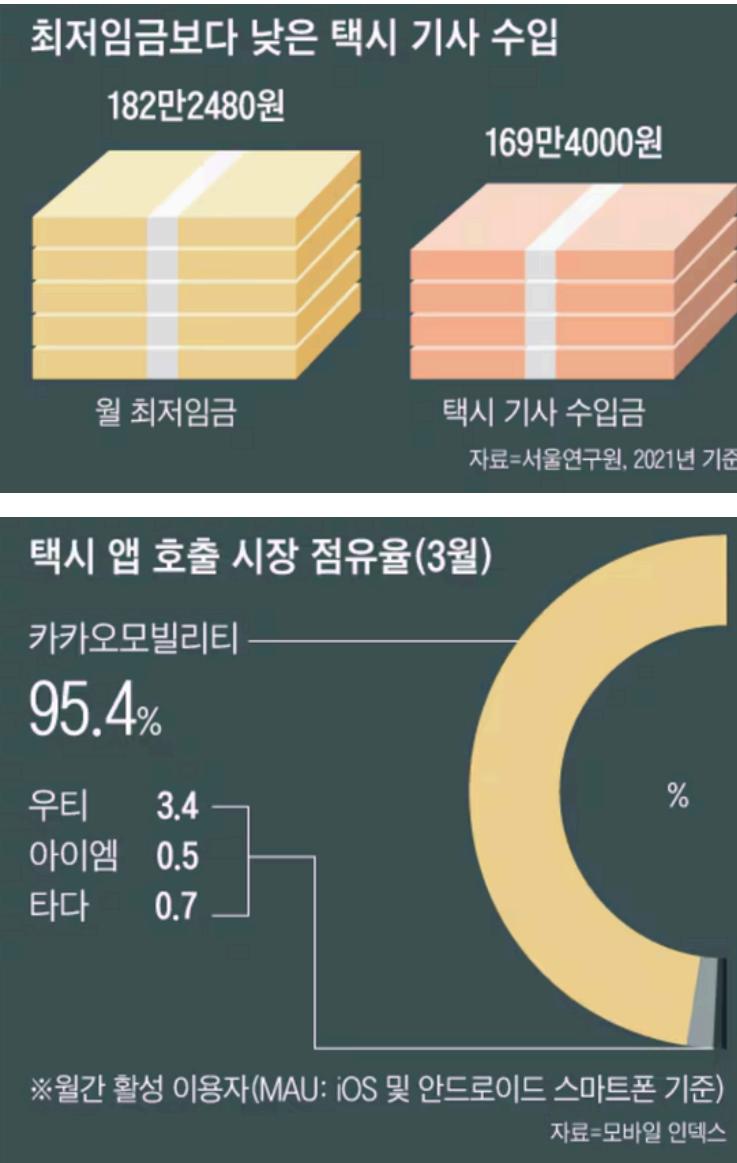
06 실험 및 평가

07 앱 화면설계

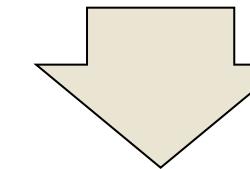
08 기대효과 및 개선방향

문제 배경 & 목적

차량 공유 서비스 및 호출 모바일 앱의 등장으로
전통적인 택시 산업은 큰 도전을 직면



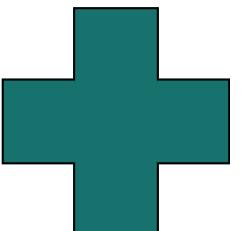
- 장시간 대기, 비효율적 운행으로 수익성 감소
- 기존 K 회사와 같은 연결 시스템에 의존해 승객을 확보, 이에 대해 수수료를 지불
- 기사들에게 경제적 부담



더 나은 수익을 창출할 수 있는
효율적인 운행 전략을
제시하는 서비스를 제공하고자 한다!

활용 데이터 및 전처리 과정

대전 시내 205대의 택시 영업 데이터
(2023년 4월부터 2024년 3월까지
1년간의 영업 내역)



기본적인 날짜 필터링 및 결측값 제거

한국 공휴일 API

공휴일 칼럼 추가

종관기상관측(ASOS)

기온, 강수량 칼럼 추가

대전광역시 택시 승강장 현황

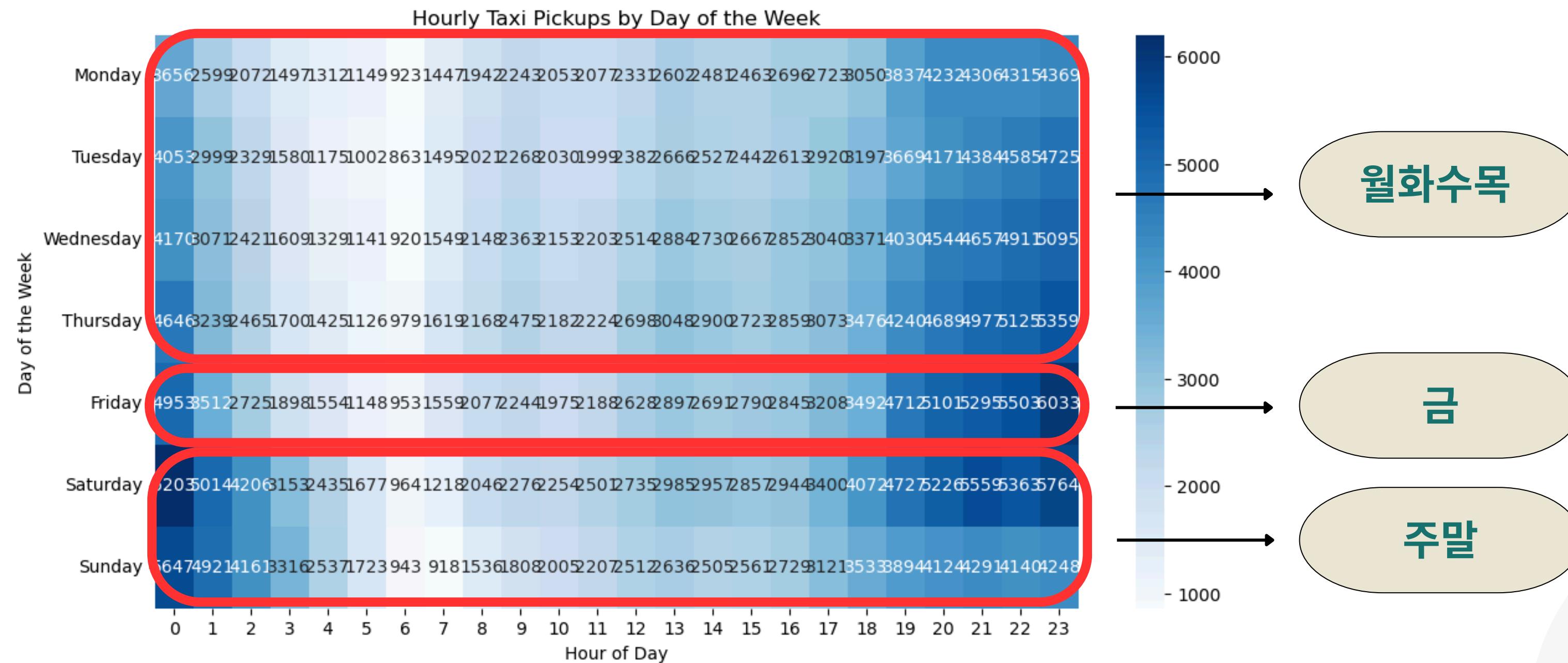
주소를 위도, 경도로 변환

대전광역시 인기 관광지 현황

주소를 위도, 경도로 변환

EDA - 히트맵

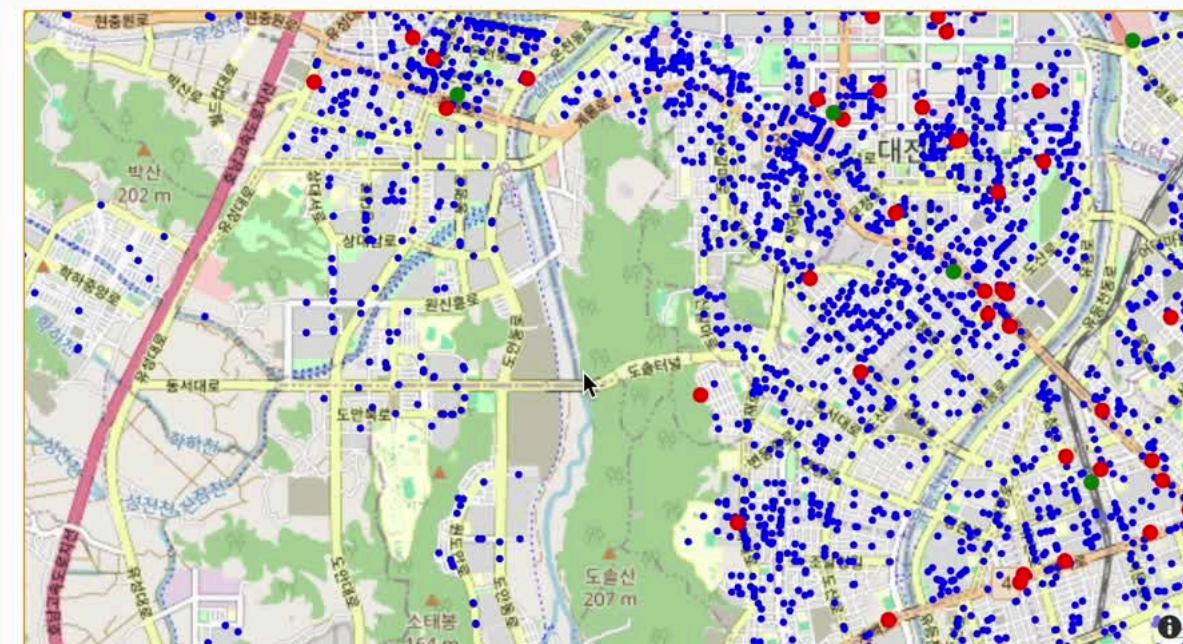
요일 시간대별 택시 수요 시각화



EDA - 지도 시각화

택시 승강장 현황, 인기 관광지 현황 데이터 추가 활용 시각화

월화수목

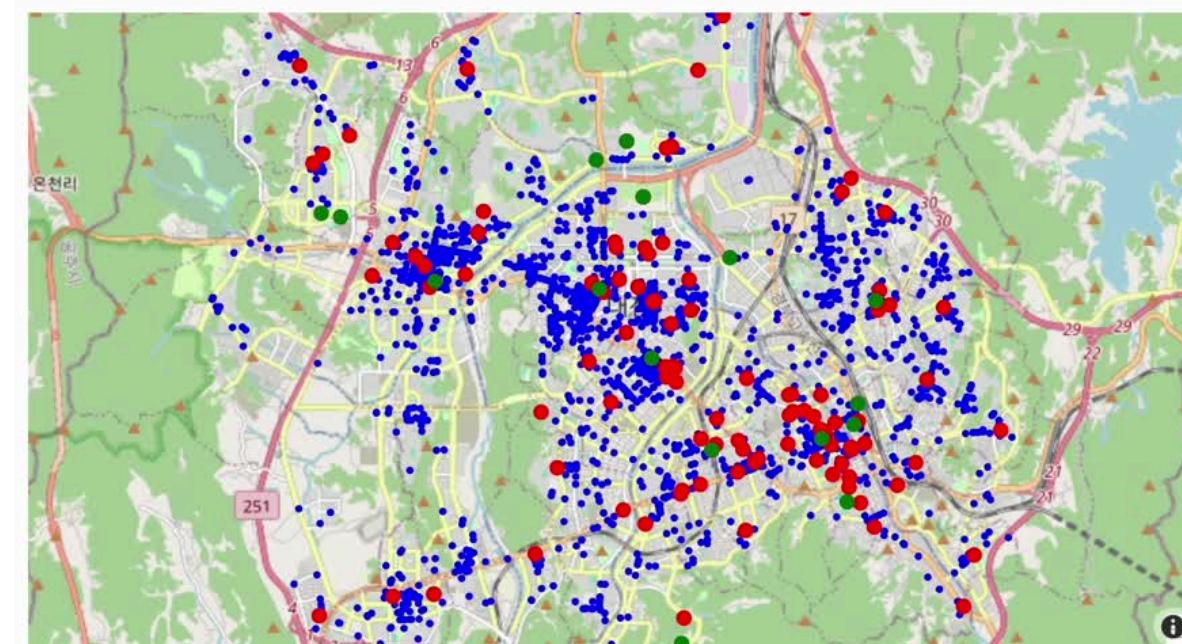


Hour: 7

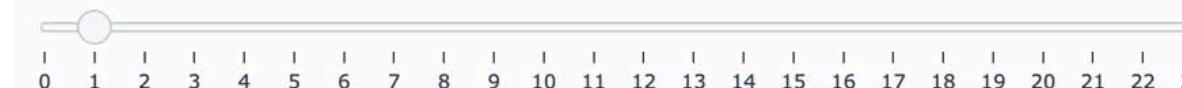


0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23

금

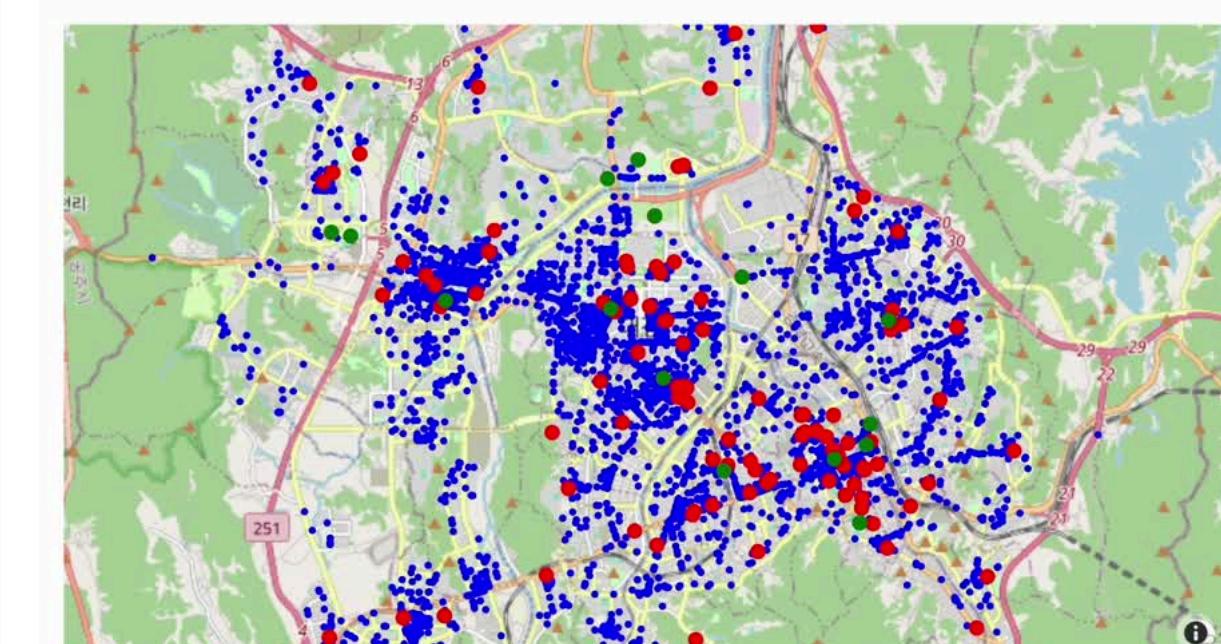


Hour: 1



0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23

주말



Hour: 3



0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23

● 택시 승차 좌표

● 택시 승강장 좌표

● 인기 관광지 좌표

클러스터링

엘보우 방법 및 실루엣 스코어로 $k=3$ 선정 후,
현재 위치, 요일, 시간, 날씨, 공휴일 유무의 특성을 반영한 **k-means**로 클러스터링 진행

주중

- 01 평일 쌀쌀한 아침 시간대
- 02 평일 비 내리는 저녁 시간대
- 03 평일 공휴일 낮 시간대

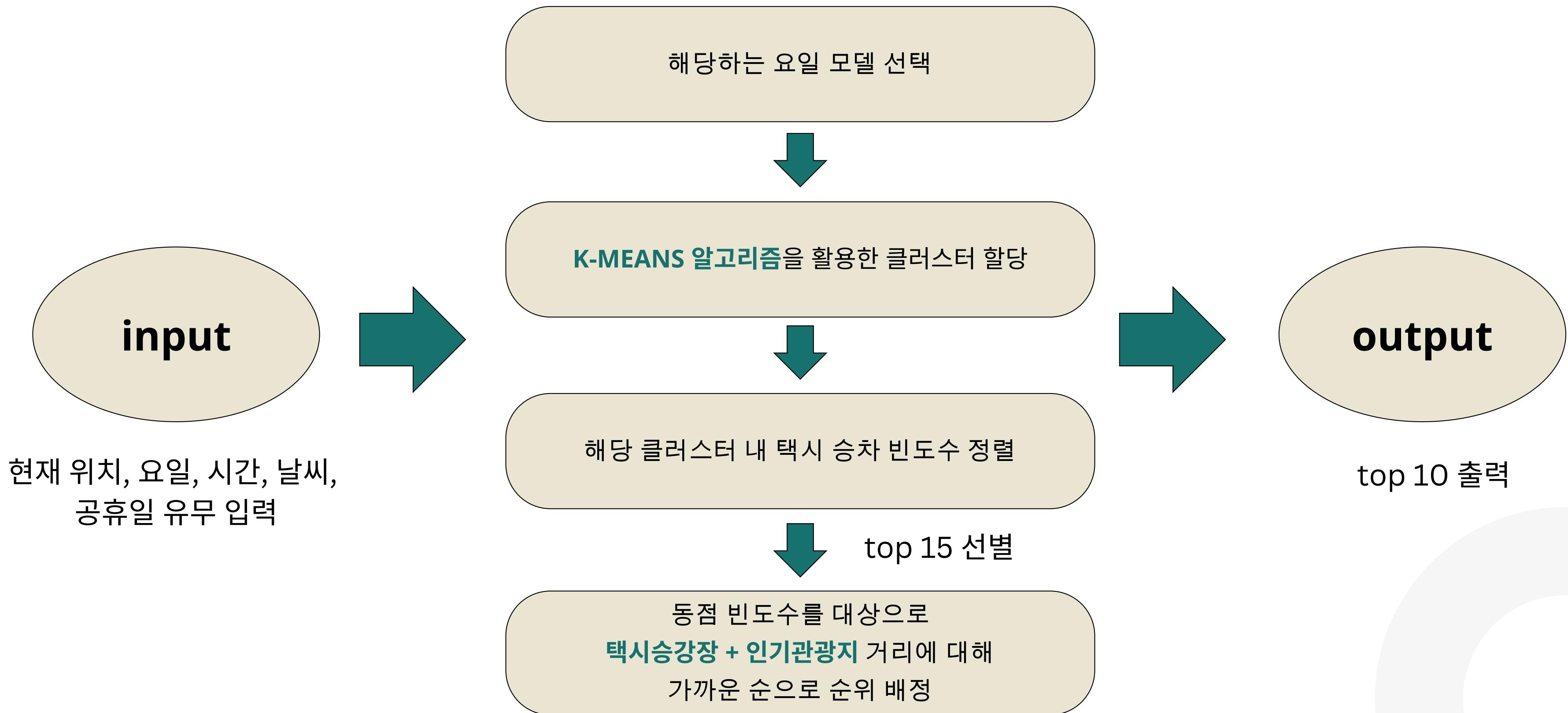
금요일

- 01 금요일 흐린 저녁 시간대
- 02 금요일 쌀쌀한 새벽 시간대
- 03 금요일 공휴일 비 내리는 정오 시간대

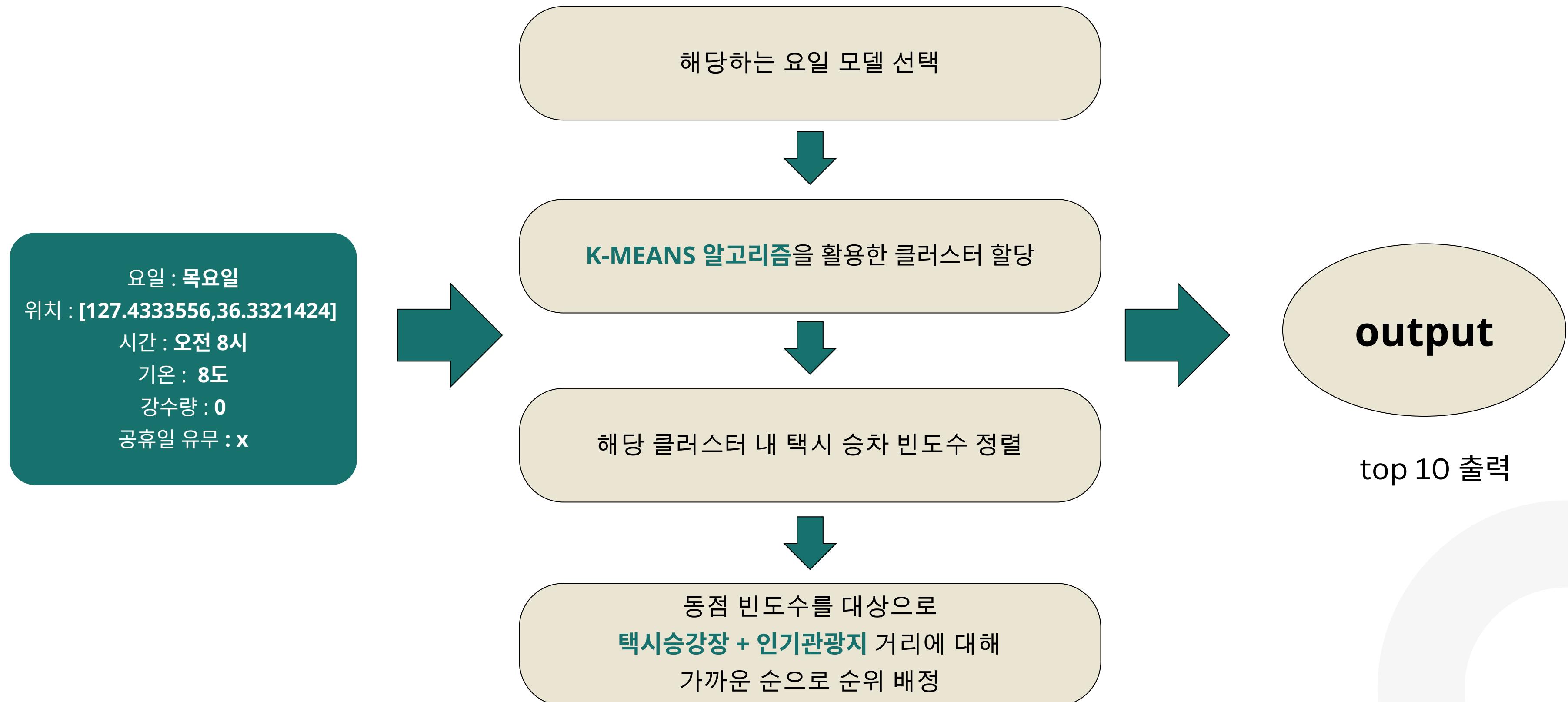
주말

- 01 주말 쌀쌀한 새벽 시간대
- 02 주말 따뜻한 저녁 시간대
- 03 주말 공휴일 추운 오후 시간대

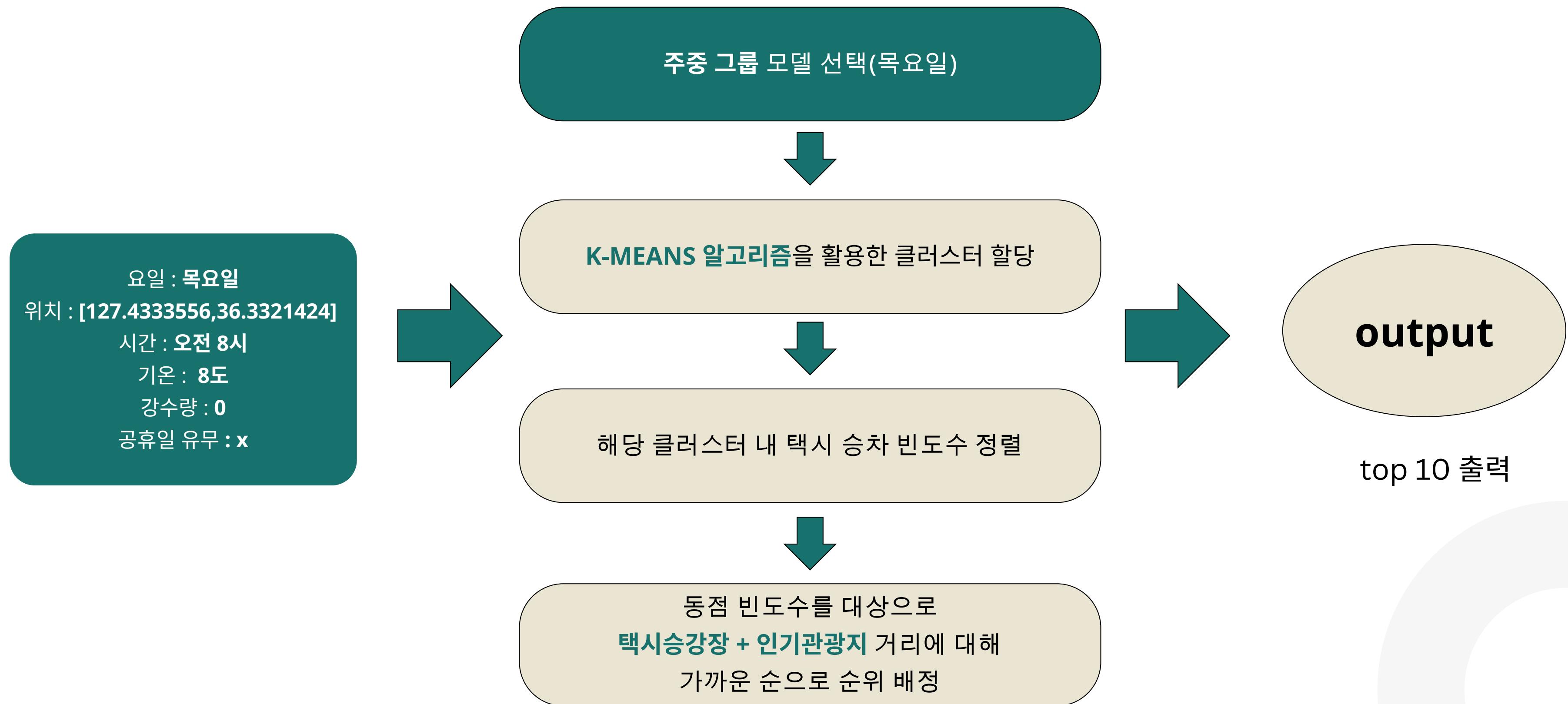
추천 모델 알고리즘



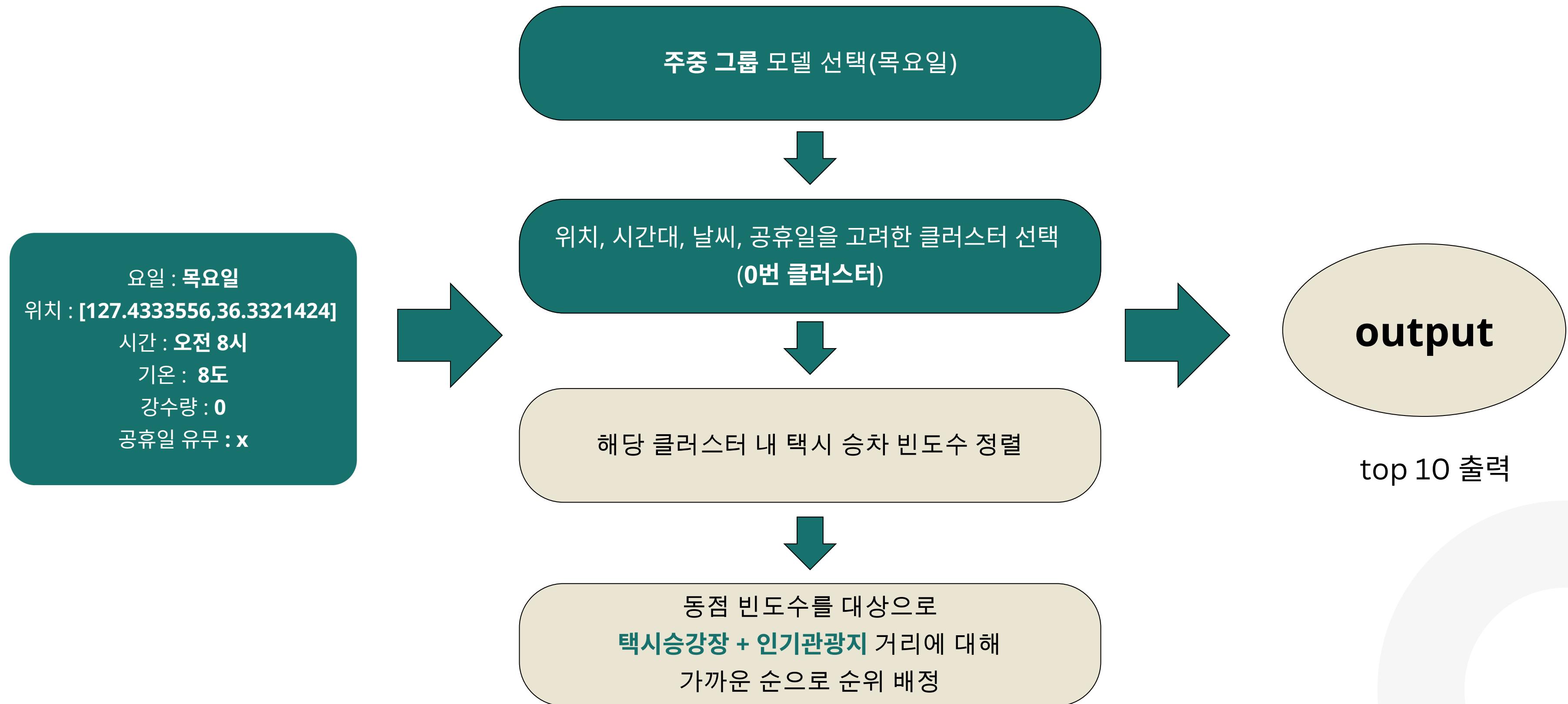
추천 모델 알고리즘



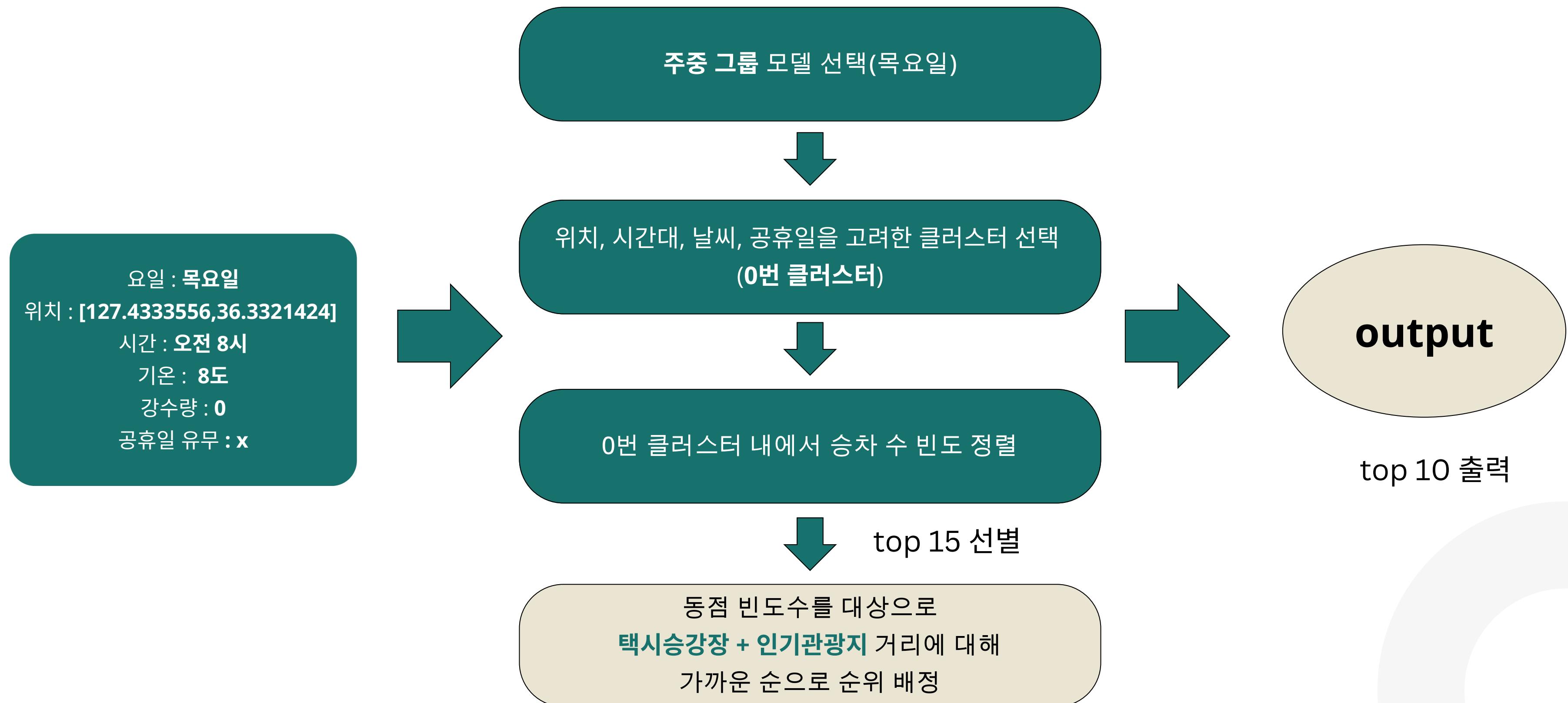
추천 모델 알고리즘



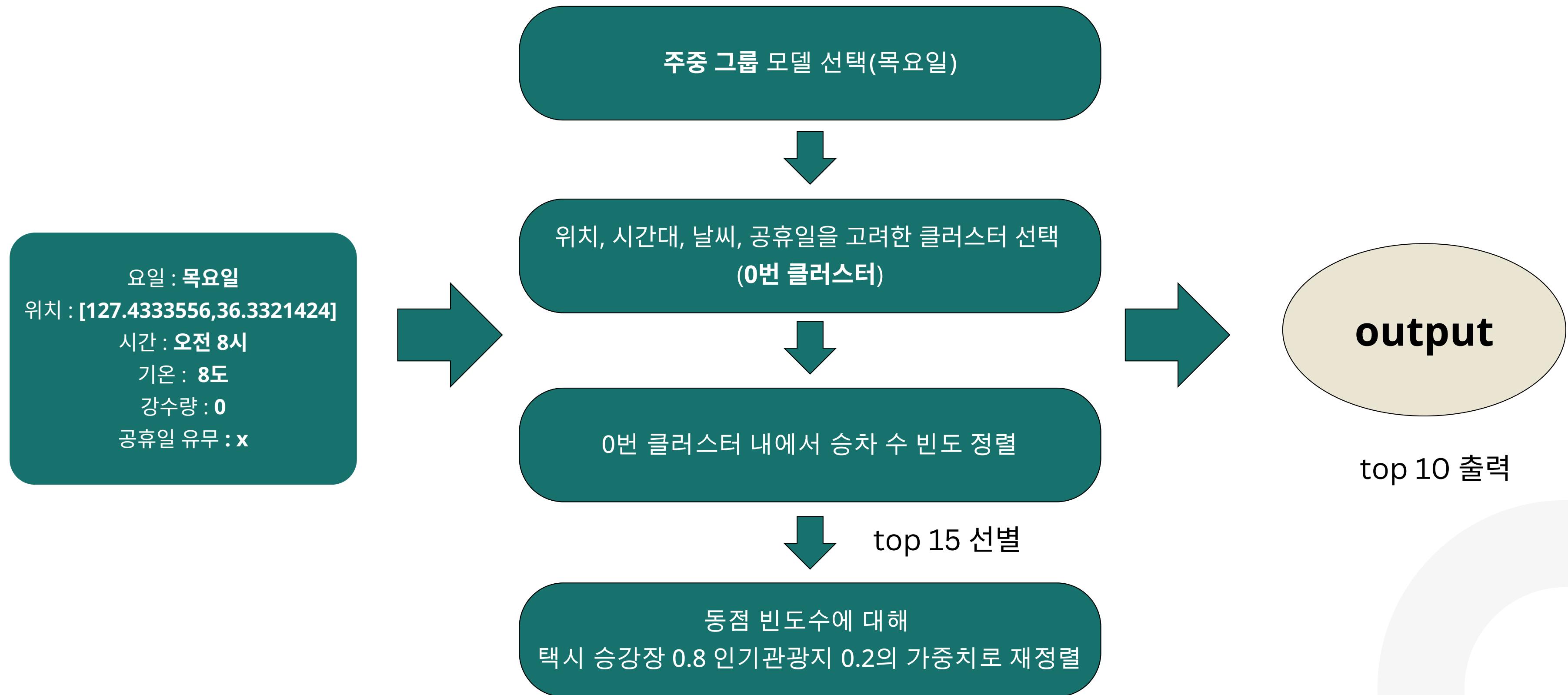
추천 모델 알고리즘



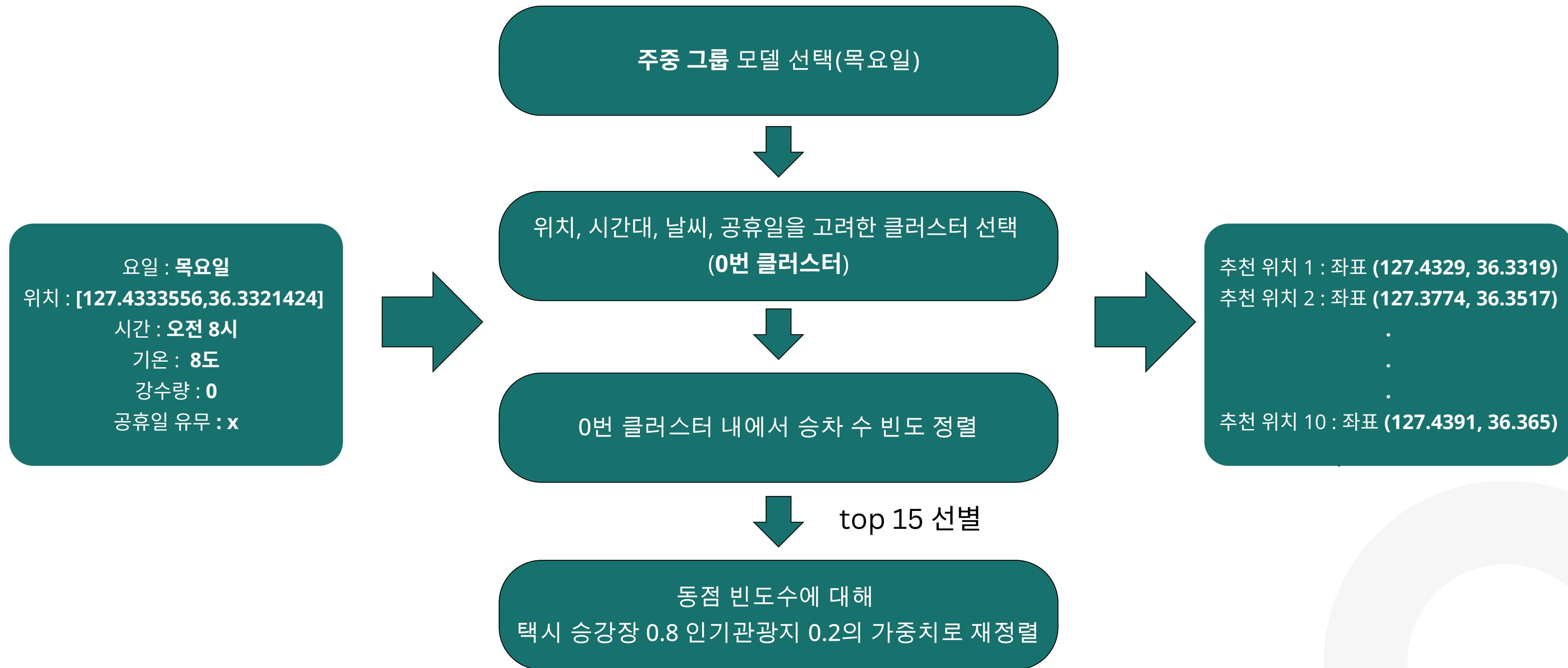
추천 모델 알고리즘



추천 모델 알고리즘



추천 모델 알고리즘



시뮬레이션

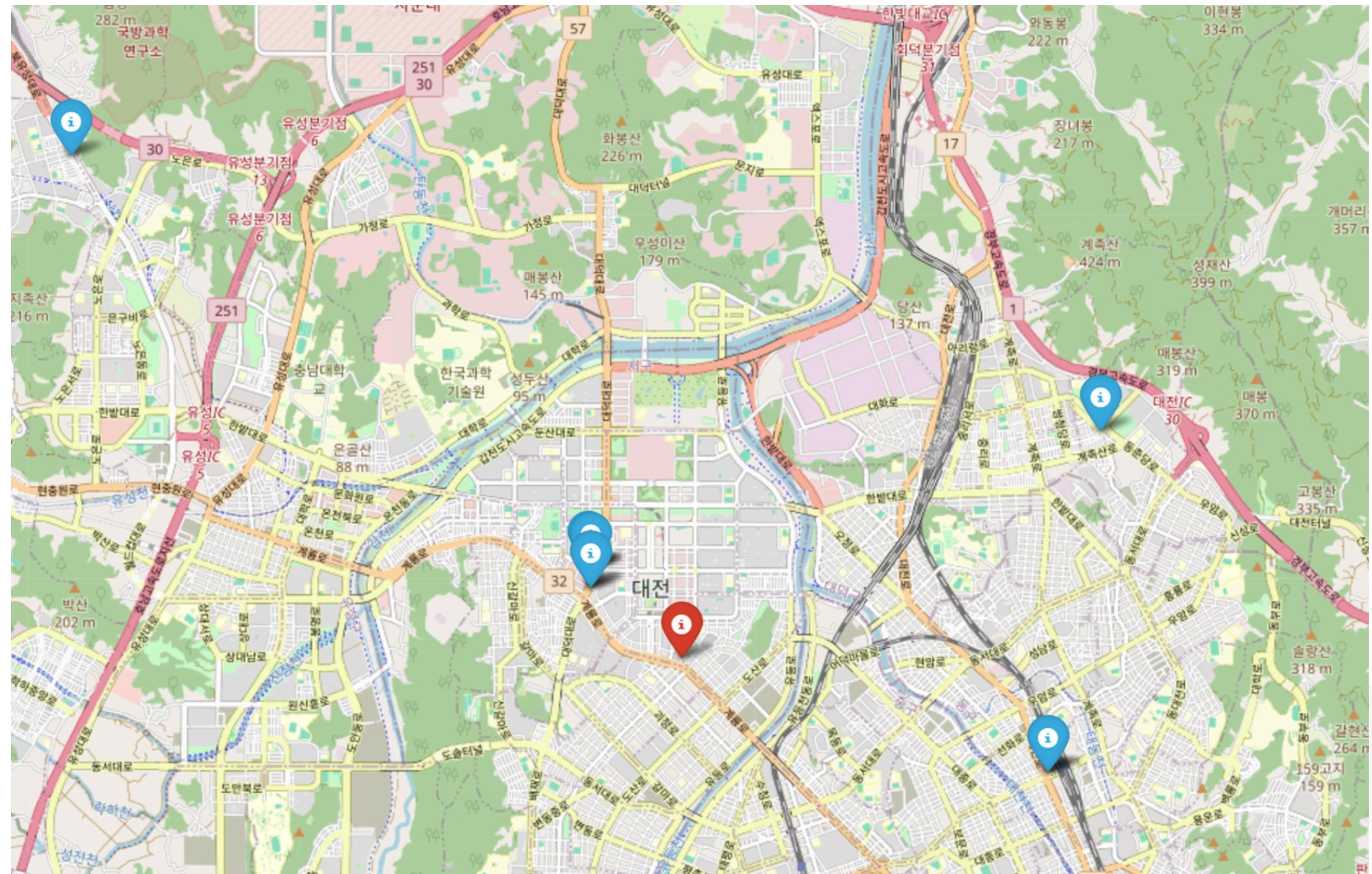
```
day = 'Thursday'
current_coords = [127.388463, 36.342952] # 예시 좌표 (경도, 위도)
current_time = 8
current_temp = 10 # 현재 기온 (예시)
current_rain = 0
is_holiday = 0

# 추천 위치 출력
top_10_locations_with_weights = recommend_locations_with_weights(
    day, current_coords, current_time, current_temp, current_rain, is_holiday,
    tourist_weight=0.8, taxi_weight=0.2
)
```

현재 상황과 유사한 승차 패턴을 추천합니다.

평일 쌀쌀한 아침 시간대 승차가 많은 곳 리스트

추천 위치 1: 좌표: (127.4329, 36.3319), 빈도: 399, 가중치 적용 점수: 0.13
추천 위치 2: 좌표: (127.3774, 36.3517), 빈도: 278, 가중치 적용 점수: 0.13
추천 위치 3: 좌표: (127.3773, 36.3517), 빈도: 276, 가중치 적용 점수: 0.13
추천 위치 4: 좌표: (127.3775, 36.3517), 빈도: 197, 가중치 적용 점수: 0.12
추천 위치 5: 좌표: (127.4328, 36.3319), 빈도: 186, 가중치 적용 점수: 0.13
추천 위치 6: 좌표: (127.3772, 36.3518), 빈도: 175, 가중치 적용 점수: 0.13
추천 위치 7: 좌표: (127.3776, 36.3499), 빈도: 139, 가중치 적용 점수: 0.28
추천 위치 8: 좌표: (127.3148, 36.3918), 빈도: 103, 가중치 적용 점수: 2.44
추천 위치 9: 좌표: (127.3775, 36.3498), 빈도: 102, 가중치 적용 점수: 0.30
추천 위치 10: 좌표: (127.4391, 36.365), 빈도: 95, 가중치 적용 점수: 1.32



● 현재 위치

● 추천 위치

실험 및 평가

검증 데이터 셋 준비

24시간

X

7일

X

20장소

= 3360 case

=> 각 case에 대한 10개의 추천 위치를 도출
(3360행) 한 후, 추천된 10개의 위치가
테스트 데이터셋의 해당 요일 그룹의 시간대에
택시 승객이 있는 경우 정답으로 간주하여
precision-recall 계산

2023년 7월 29일 토요일 새벽 2시
추천 좌표 10개 정답률을 확인하려면..?

테스트 데이터에서 토요일에 해당하는
주말 그룹의 2시 승차 좌표들 중 추천된 10개의
좌표가 몇 개나 포함되는지 확인

70% 정답률 : 40 case

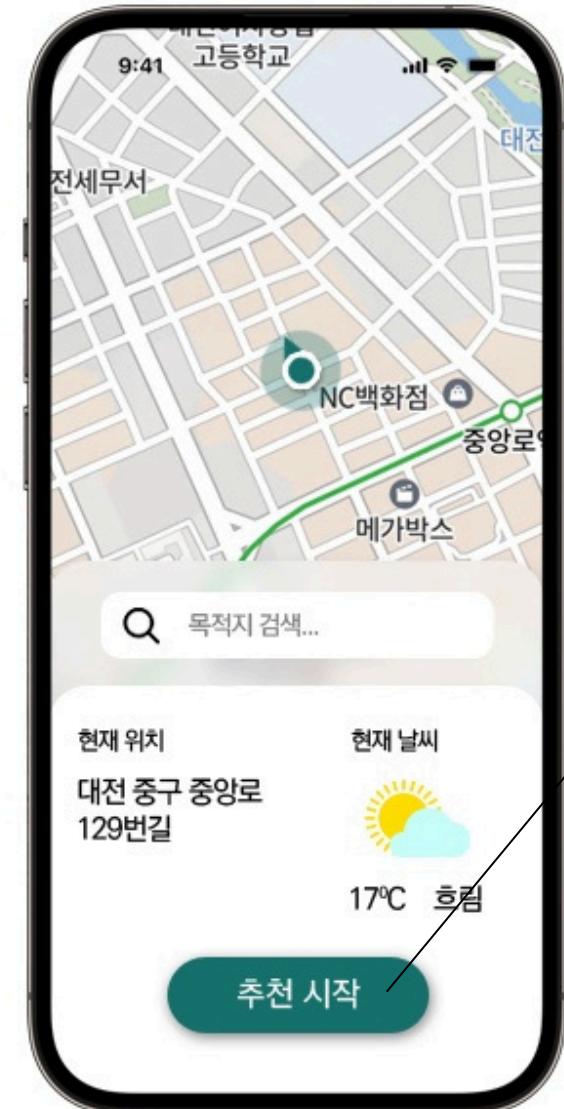
60% 정답률 : 240 case

•
•
•

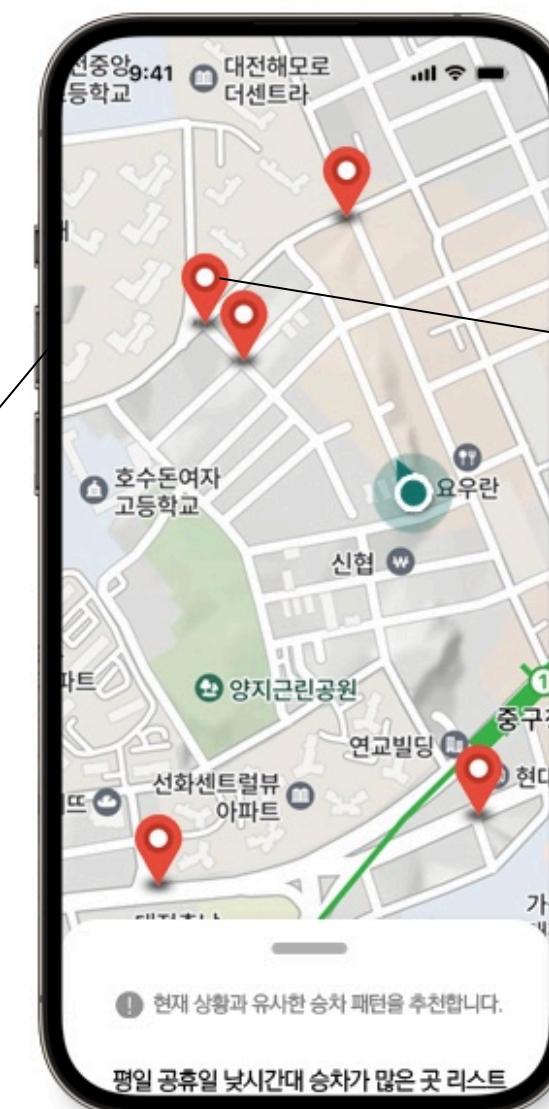
앱 화면 설계



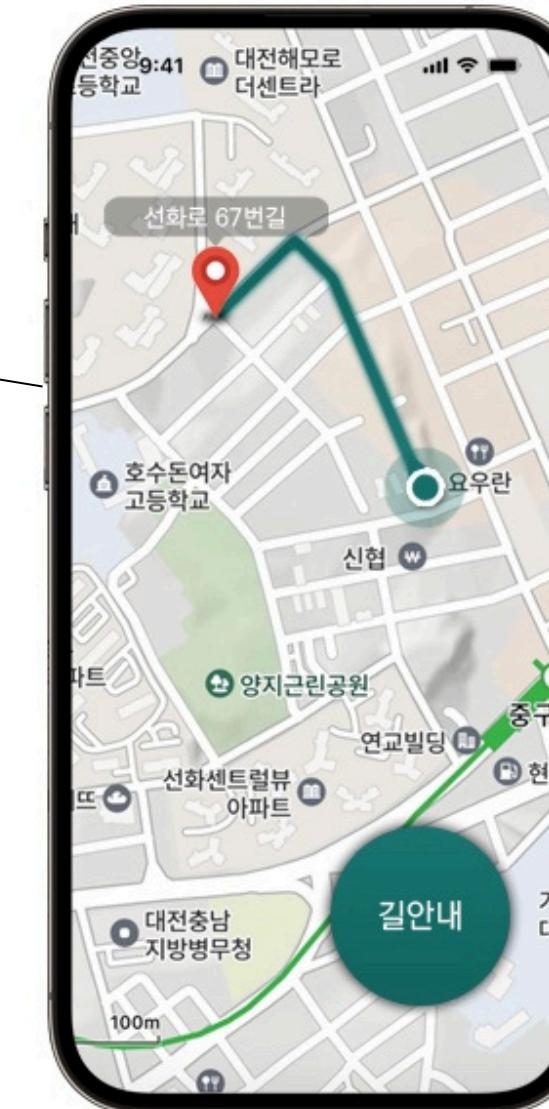
<GO! 캡틴 로고>



<앱의 시작화면>



<추천 승차장소 화면>



<특정 장소 선택 후 경로 안내 화면>

개선 방향

알고리즘 구조

```
weighted_score = (min_tourist_distance *  
tourist_weight) + (min_taxi_distance * taxi_weight)  
+ (새로운 변수1) * (가중치) + (새로운 변수2) * (가중치) + α
```

변수의 중요도에 따라 가중치를 반영하여 위치를 추천!

다른 변수들을 추가로 발전시킬 수 있는 가능성을 확인

기사 성향 클러스터링

엘보우 방법 및 실루엣 스코어로 k=3 선정 후 KMeans로 클러스터링 진행.
=> 승차 거리, 요금, 할증여부, 주행시간 반영.

- 01 단거리 운행 패턴
- 02 장거리 운행 및 고요금 지역에서의 운행 패턴
- 03 할증 운행을 자주 수행하면서도 비교적 짧은 거리의 운행 패턴

Feature	Importance
할증여부	0.382649
요금	0.286235
승차거리(㎞)	0.183617
주행시간_초	0.147499

변수 중요도 파악

기대 효과

택시 기사

택시 기사 수익 증대

불필요한 공차 시간을 줄이고 높은 수익을 창출

승객

승객들의 만족도 제고

택시가 빠르게 도착함으로써 승객들의 대기 시간 단축

대전시

교통 흐름 개선

다양한 변수를 고려한 최적화된 이동 경로를 제시함으로써 택시 운송 효율 향상



2024 DATA AI 분석 경진대회



THANK YOU

November 2024