

# 추천 시스템을 활용한 새로운 영업 패러다임 제안

빅데이터 분석에 기반한 맞춤형 서비스

- 경기대학교 홀인원 팀 -

김재희 손현우 신다은 이세호 최세은 홍연정



# CONTENTS

01

프로젝트 배경

02

활용 데이터  
및 주제 정의

03

회원권  
추천 시스템

04

국내 골프장  
추천 시스템

05

분석 결과

# 01 | 프로젝트 배경

## 골프 트렌드 탐색

### ( 골프 관련 뉴스 기사 )

#### [트렌드쑈인] "이 재밌는 걸 어른들만 했던 말이야?" 골프 시작하는 MZ세대

코로나19 사태가 장기화되면서 새로운 트렌드가 등장했다. 바로 골프다. 기존의 골프는 스포츠이기는 하지만 금전적인 여유가 있는 중장년층의 전유물로 여겨졌다.

하지만 최근 골프가 대중화 되고 있다. **MZ세대가 빠르게 골프 문화에 유입**되고 있는데, 이를 증명하듯 TV를 틀 때마다 전에 없던 골프 프로그램들이 쏟아져 나온다.

실제 2010년대 매년 4~5조원에 머물렀던 국내 골프장 매출이 지난해 7조원대로 성장했다. 그 중에서도 MZ세대 골프 수요가 지난해 100만 명을 넘어섰다. KB금융지주 경영연구소에 따르면 이른바 **골린이**로 불리는 3년 이하 골프 입문자 중 2040대가 65%를 차지했다.

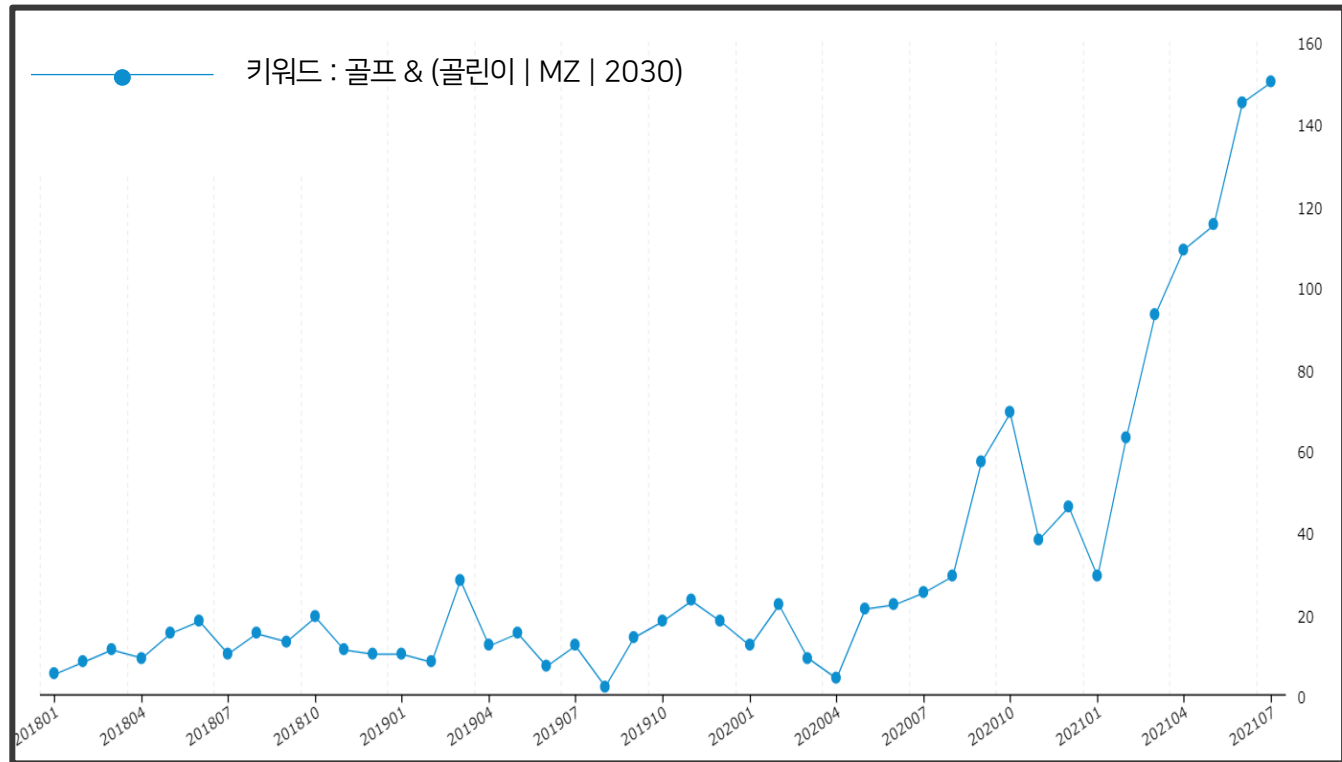
엠브레인 트렌드모니터의 골프 인식 조사에 따르면 전체 응답자의 73.1%가 예전보다 젊은 층의 골프 인구가 많아진 것 같다고 느끼고 있었으며, 요즘 골프는 젊은이들도 많이 즐기는 운동이라고 바라보는 시각이 66.4%에 달했다.

MZ세대 골프 인구가 증가함에 따라 **골프 관련 소비가 늘어나면서 골프 의류, 용품, 카드사, 골프 예능까지** 골프와 직간접적으로 관련 있는 산업군에서 젊은 골퍼를 잡기 위한 마케팅이 한창이다.

“골프에 관심을 갖는 **연령대가 다양**해짐

### ( 골프 관련 뉴스 건수 변화 )

기간 : 2018-01-01 ~ 2021-07-31  
출처 : 빅카인즈(BIG KINDS)

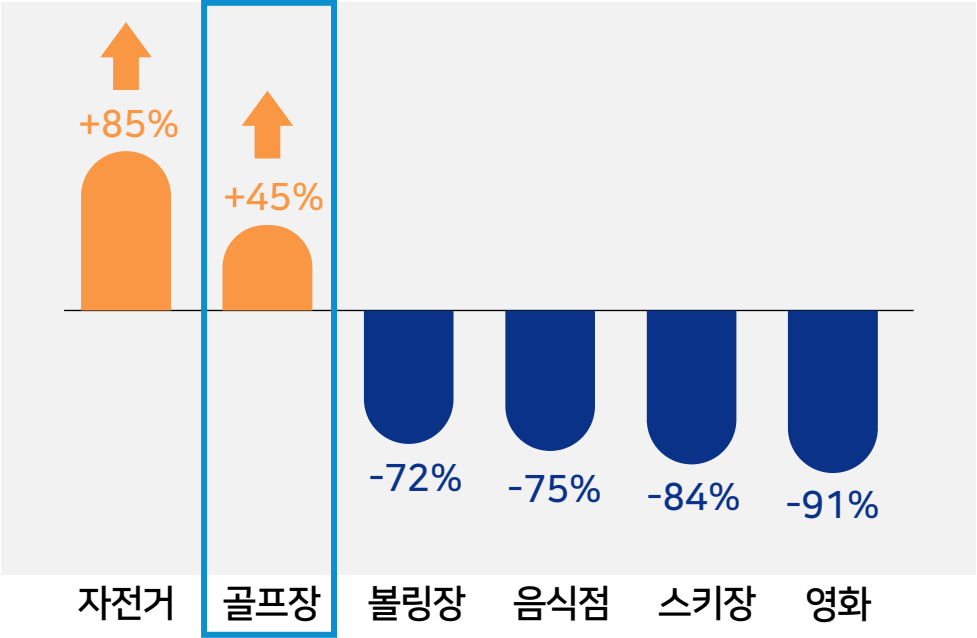


2018년 1월 이후 골프 관련 뉴스 건수 **증가**

골프 트렌드 탐색

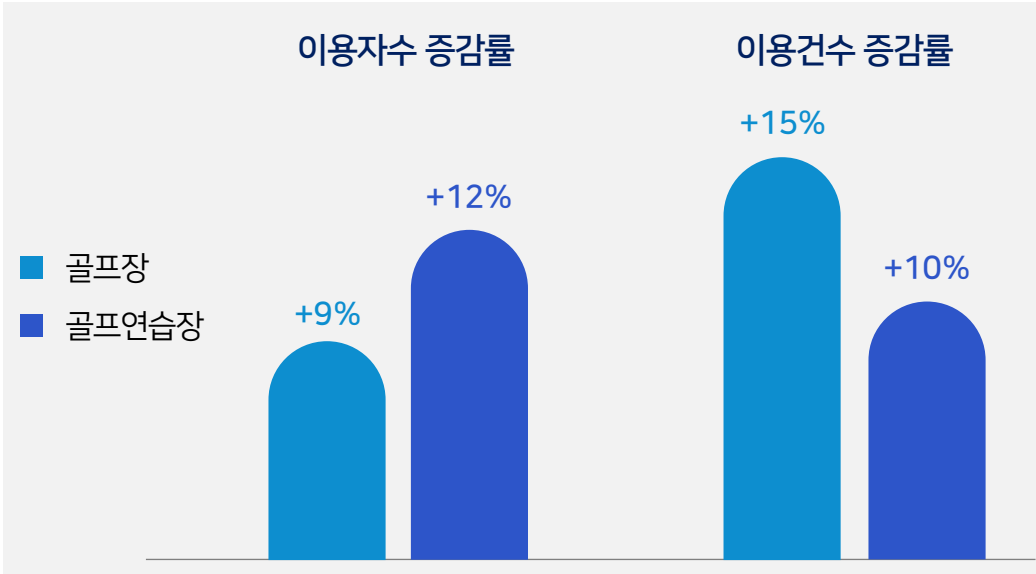
삼성카드 결제 데이터  
(기간 : 2019년 1월 ~ 2021년 4월)

코로나 이전 대비, 여가 활동 관련 카드 이용금액 증감률



\* 분석기준: 코로나19 이전(2019년 1~4월) / 코로나19 이후(2021년 1~4월)의 카드 이용금액

골프장 및 골프연습장 이용 관련 추이



\* 분석기준: 코로나19 이전(2019년 1~4월) / 코로나19 이후(2021년 1~4월)의 골프장 및 골프연습장 이용자수

“ 코로나 19 유행에도 골프장/골프 연습장의 이용이 증가 ”

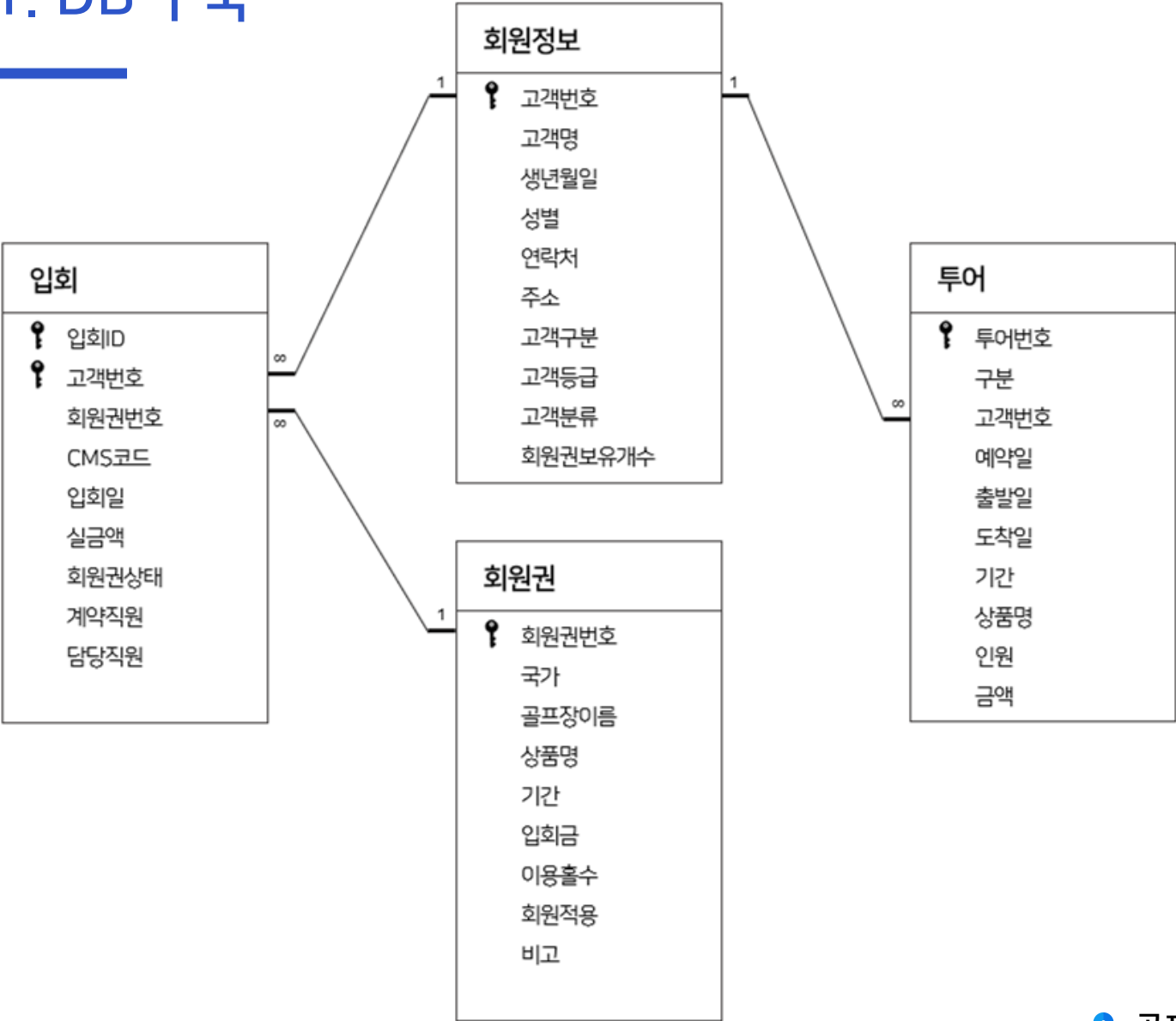
# 02

## 활용 데이터 및 주제 정의

01. DB 구축

02. 주제 정의

01. DB 구축



회원정보 테이블

입회하지 않은 상담 고객까지 포함한  
모든 고객에 대한 정보

회원권 테이블

판매 중단된 회원권까지 포함한  
현재까지의 모든 회원권 정보

입회 테이블

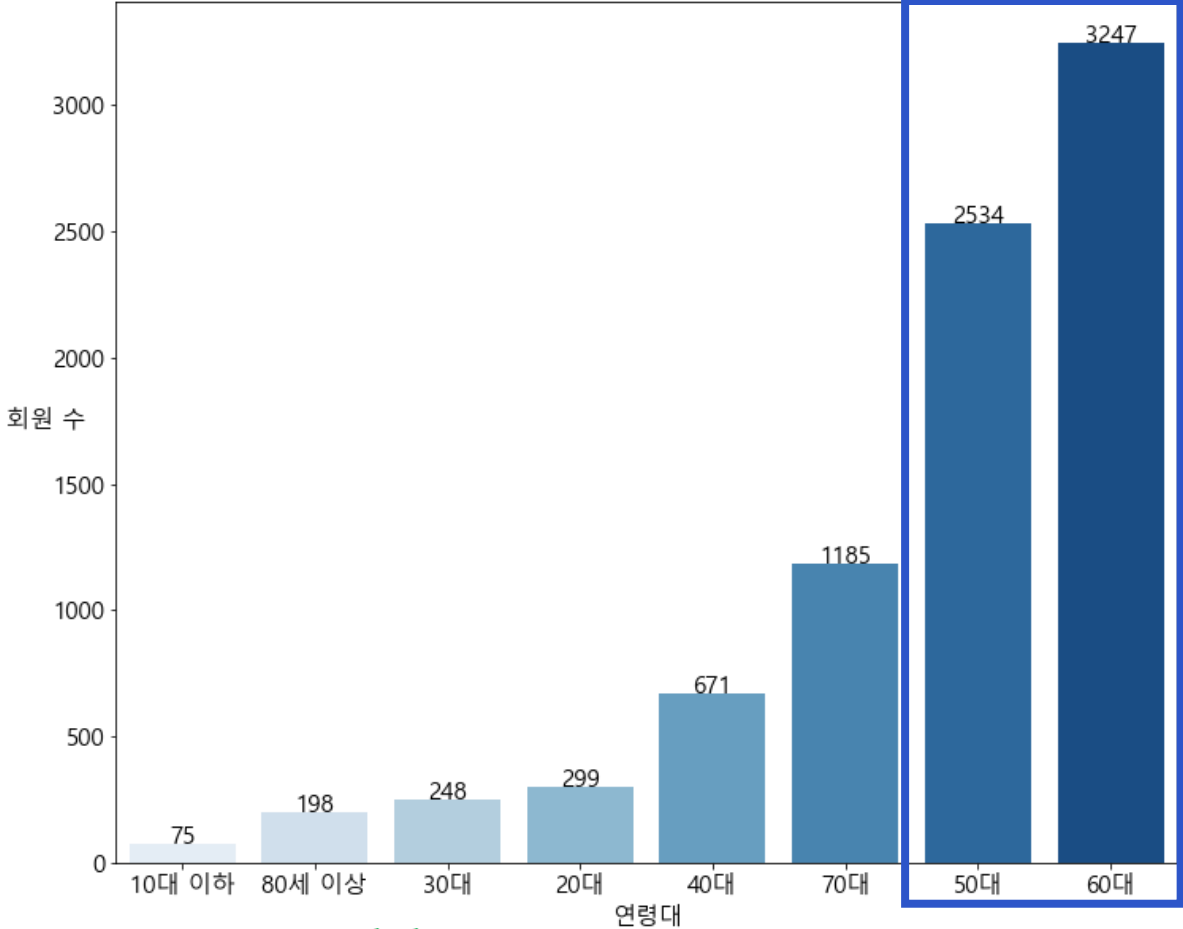
회원권을 보유하고 있는 고객에 대한 입회 정보

투어 테이블

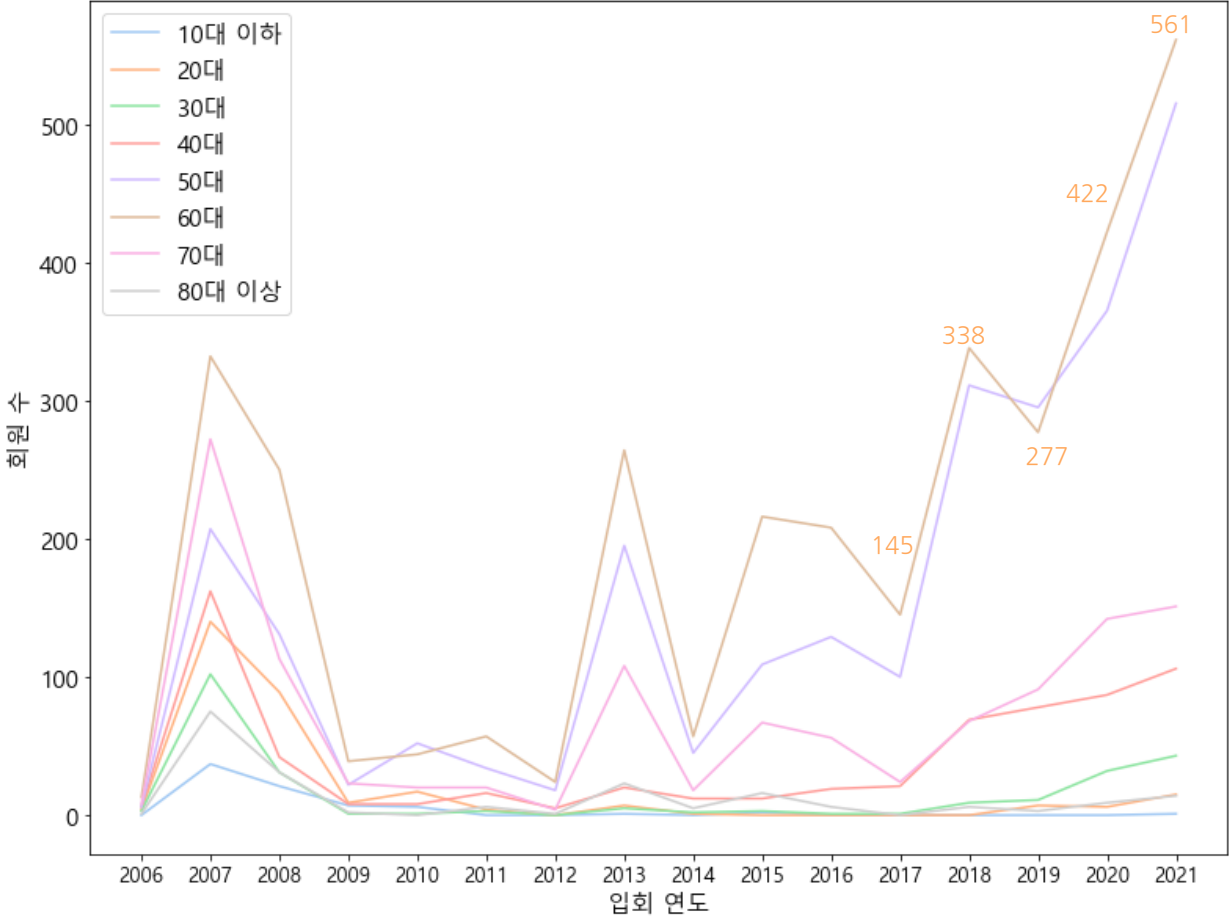
국내 및 해외 투어 관련 정보

## 02. 주제 정의

연령대별 입회 건수



연도별 및 연령대별 입회 회원 수



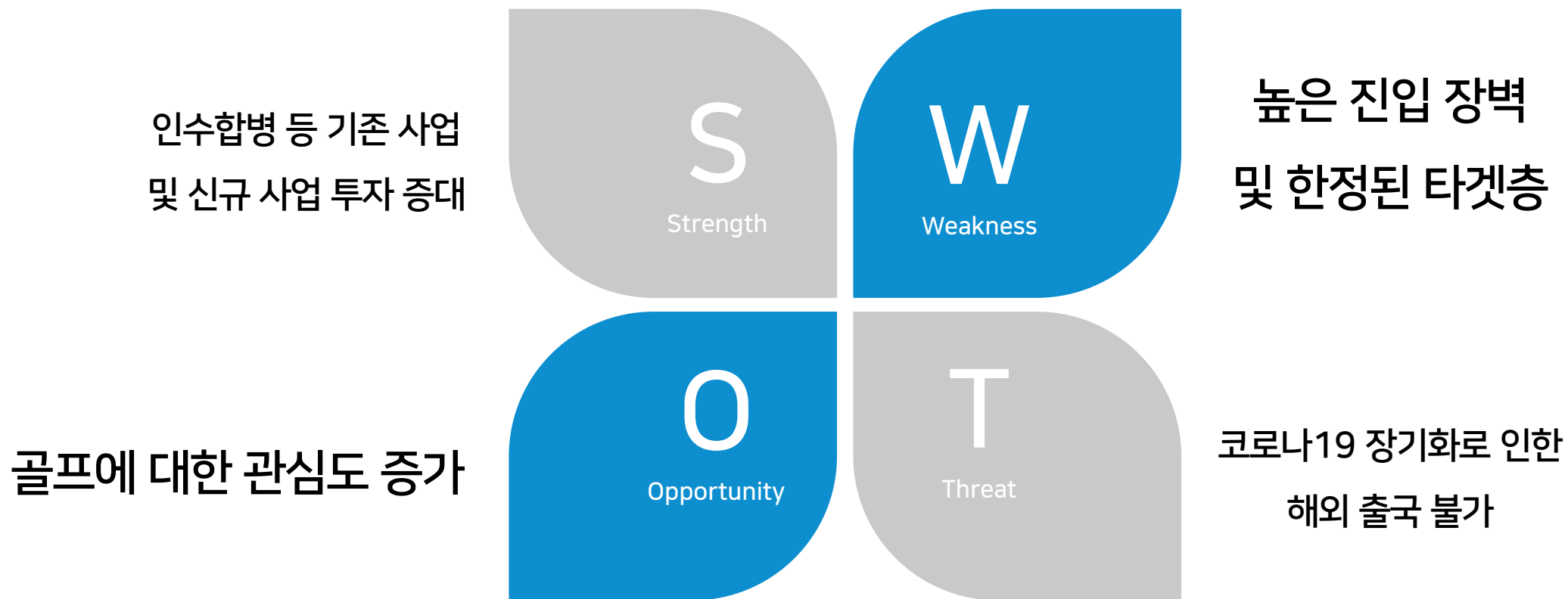
“

한케이골프의 주 고객층은 50,60대이며, 모든 연령대에서 높은 성장세를 보이고 있음

”

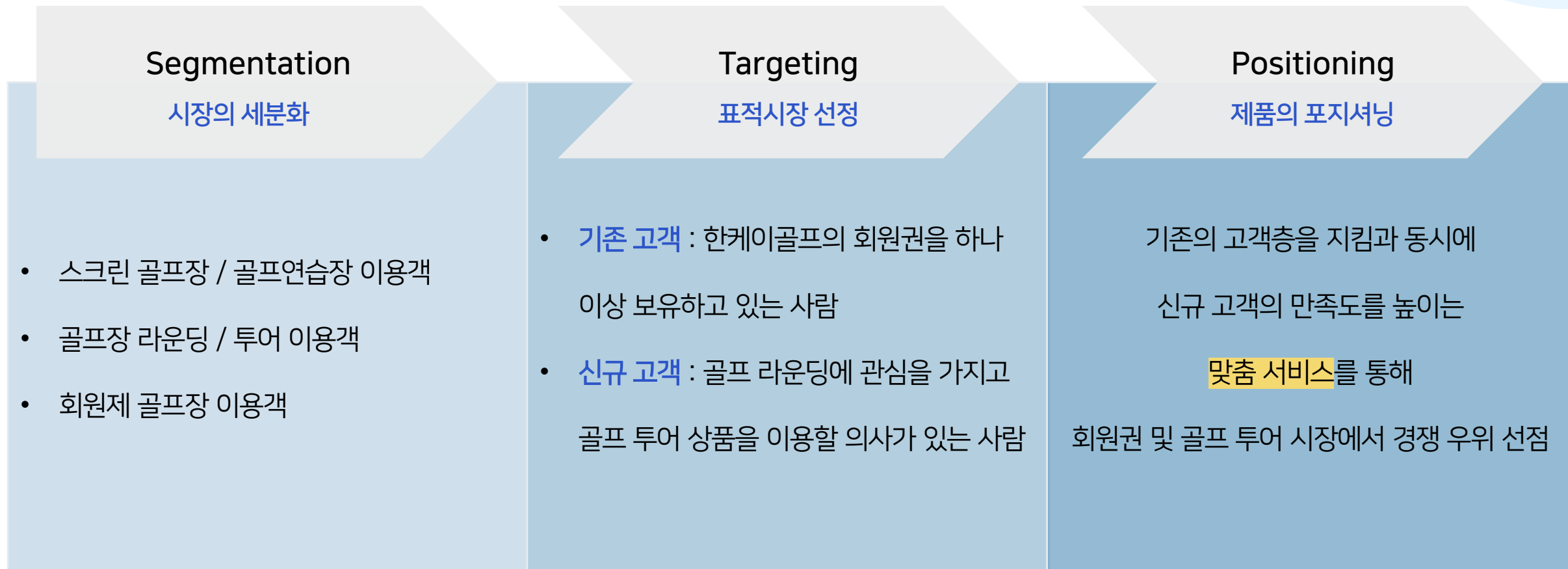


## 02. 주제 정의 - SWOT 분석



사회 전반적인 골프관심도의 증가와 더불어 **현재의 캐시카우 사업을 영위하면서**  
**새롭게 유입될 가능성이 있는 고객군을 위한 사업 확장의 필요성 대두**

## 02. 주제 정의 - STP전략



“

기존/신규 고객의 니즈를 모두 충족시킬 수 있는 **고객 맞춤형 서비스 제안**

”

# 03

## 회원권 추천 시스템

01. 프로세스

02. 한계점

01. 프로세스 - 회원권 기본 정보



회원권 기본 정보



추가 데이터 수집



크롤링 데이터 전처리



회원권 유사도 계산



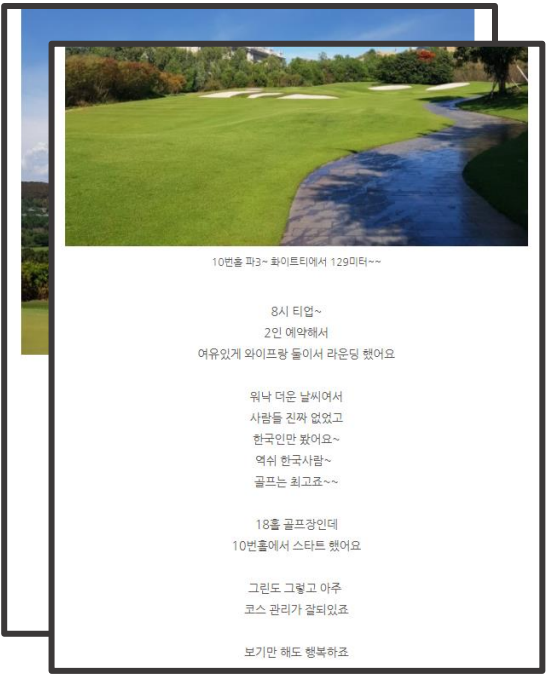
회원권 추천

회원권번호	국가	골프장이름	상품명	기간	입회금	이용률수	회원적용	비고
0	CB-01	중국	백작36	가족	마감	28000000	36.0	부부+18세이하 자녀
1	CB-02	중국	백작36	법인		50000000	36.0	
2	CC-01	중국	청도18	법인	마감	45000000	18.0	기명1+무기명1
3	CC-02	중국	청도18	개인		19800000	18.0	
4	CC-03	중국	청도18	가족		24800000	18.0	
...	...	...	...	...	...	...	...	...
65	VV-02	베트남	빈펄	GOLD	20년	19200000	90.0	기명1+무기명3 남호이안, 푸꾸옥, 하이퐁, 나트랑 4곳 모두 이용
66	VV-03	베트남	빈펄	SILVER	20년	14400000	90.0	기명1+무기명3 남호이안, 푸꾸옥, 하이퐁 중 1곳만 이용
67	VV-04	베트남	빈펄	VIP 법인		36000000		
68	VY-01	베트남	옌중	VIP	평생	14000000	36.0	무기명4
69	VY-02	베트남	옌중	GOLD	평생	9000000	36.0	기명1+무기명3

- 회원권번호 : 회원권 구별을 위한 임의의 코드
- 상품명 : 회원권 상품명(가족/법인 , VIP/VVIP 등)
- 기간 : 회원권 보장 기간
- 회원적용 : 회원권 적용 범위

구축한 데이터베이스의 '회원권' 테이블

01. 프로세스 - 추가 데이터 수집



일일체류비	그린피	카트비	캐디피	숙소	조식	중식	석식	식사	캐디팁	보험비	라커룸	송영	dist
540000	1	1	1	1	1	0	0	0	1	0	0	1	2991.006627
417000	1	1	1	0	0	0	0	1	1	0	0	0	2656.037250
276000	1	1	0	1	1	1	1	3	0	0	0	0	4575.061433
399200	1	1	1	1	1	0	0	1	1	0	0	0	2651.114670
540000	1	1	1	1	1	0	0	1	0	1	1	0	2019.247427

\* dist는 인천공항과 각 골프장 간의 Haversine Formula를 활용한 최단 거리로 계산

01 블로그 게시물 크롤링

02 일일체류비와 골프장 거리

(일일체류비의 서비스 포함 여부는 포함(1) / 미포함(0)으로 표현)

## 01. 프로세스 - 블로그 크롤링 데이터 전처리



회원권 기본 정보



추가 데이터 수집



크롤링 데이터 전처리



회원권 유사도 계산



회원권 추천

```
stopwords = ['골프', '회원', '골프장', '베트남', '빈펄', '케이', '블루', '퍼시픽',
              '오이타', '랑스', '레이테',
              '입회', '평생', '가장', '한국', '일본', '무기명', '스프링', '데일',
              '정도', '포함', '하우스', '필리핀',
              '지역', '문의', '골퍼', '가능', '바로', '하노이', '현대', '때문',
              '세부', '준비', '시간', '치링', '혜택', '설계', '아주',
              '라이스', '말레이시아', '대한민국', '중국', '프랑스', '비용', '오늘',
              '제주도', '남호', '이안', '하이퐁', '나트랑', '랜드']
```

불용어 지정

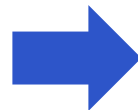


```
rst = []
tfidf = TfidfVectorizer(max_features=500, stop_words=stopwords, tokenizer=tokenizer)

for i in range(17):
    tdm = tfidf.fit_transform(result[i])
    word_count = pd.DataFrame({
        '단어': tfidf.get_feature_names(),
        '빈도': tdm.sum(axis=0).flat
    })
    word_count = word_count.sort_values('빈도', ascending=False)
    lst = list(word_count['단어'])
    rst.append(lst)
```

TF-IDF 벡터

→ 가중치(중요도) 높은 단어 추출



	nation	name	keyword
0	베트남	빈펄	[리조트, 호텔, 최고, 시설, 다낭, 바다, 아름다운, 거리, 사파리, 럭셔리, ...]
1	베트남	치링	[유산, 리조트, 호텔, 박닌, 최고, 그린, 하롱베이, 거리, 시설, 공항, 관광, ...]
2	베트남	옌중	[최고, 리조트, 유산, 호텔, 거리, 하롱베이, 도전, 자연, 시내, 관광, 공항, ...]
3	말레이시아	닐라이스프링스	[리조트, 그린, 라이, 시설, 망고, 공항, 호텔, 거리, 웨이, 수영장, 페어, ...]
4	필리핀	레이테	[리조트, 자연, 국립공원, 최고, 공항, 그린, 캐디, 장박, 추천, 다양한, 오...
5	대한민국	현대더링스	[자연, 그린, 웨이, 페어, 로얄, 링크스, 바다, 병커, 갈대, 좋은, 해안, ...]
6	대한민국	제피로스	[그린필드, 그린, 호텔, 자연, 잔디, 좋은, 플레이, 상태, 조식, 많은, 공항, ...]
7	대한민국	스프링데일	[그린, 자연, 한라산, 바다, 좋은, 편백나무, 사진, 맛집, 휴식, 난이도, 편...
8	일본	퍼시픽블루	[리조트, 온천, 공항, 바다, 최고, 가든, 거리, 그린, 조식, 석식, 아름다운, ...]
9	일본	그린랜드	[호텔, 놀이, 공원, 리조트, 라멘, 단체, 숙박, 입장권, 관광, 놀이기구, 자...
10	대한민국	오너스	[그린, 티샷, 좋은, 웨이, 페어, 해저드, 잔디, 자연, 프로, 사진, 병커, ...]
11	중국	남산	[호텔, 국제, 공항, 거리, 그린, 식당, 레슨, 시설, 최고, 좋은, 추천, 황...
12	중국	미션힐스	[리조트, 최고, 시설, 최대, 온천, 스파, 호텔, 다양한, 대회, 아시아, 규모...
13	대한민국	샤인빌	[샤인, 그린, 파크, 자연, 호텔, 야자수, 리버, 국제, 풍경, 나무, 페어, ...]
14	대한민국	용인	[용인, 그린, 블루원, 거리, 사랑, 웨이, 잔디, 페어, 코리아, 드라이버, 병...
15	대한민국	이븐데일	[그린, 페어, 웨이, 티샷, 사진, 드라이버, 잔디, 청주, 병커, 난이도, 거리...
16	대한민국	타미우스	[미우스, 플레이, 자연, 회장, 대회, 그린, 오름, 비즈, 아름다운, 휴식, 잔...

골프장별 주요 키워드 추출

## 01. 프로세스 - 회원권 유사도 계산



회원권 기본 정보



추가 데이터 수집



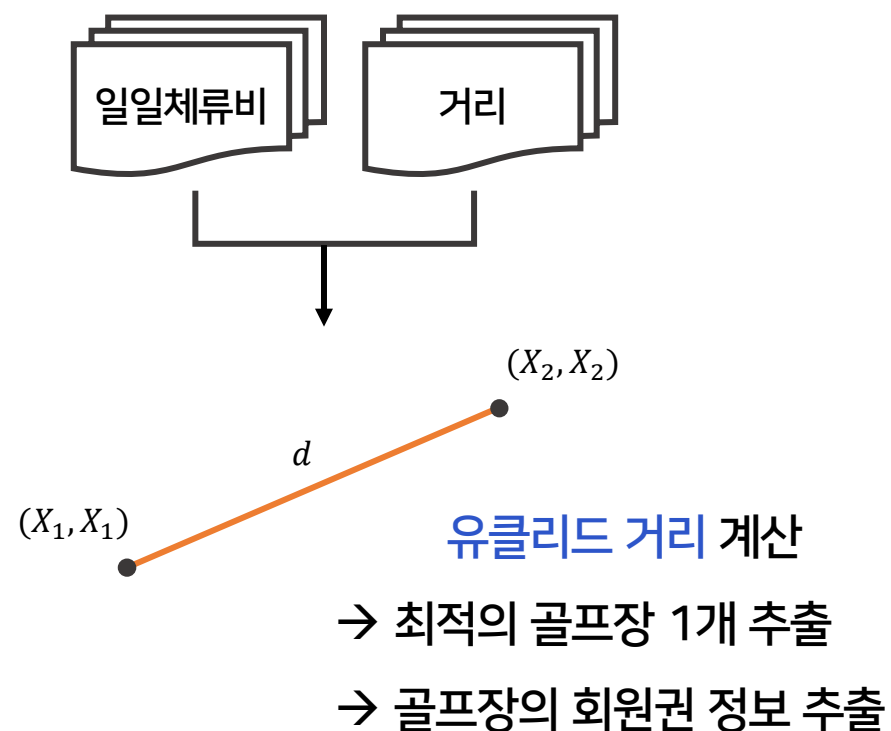
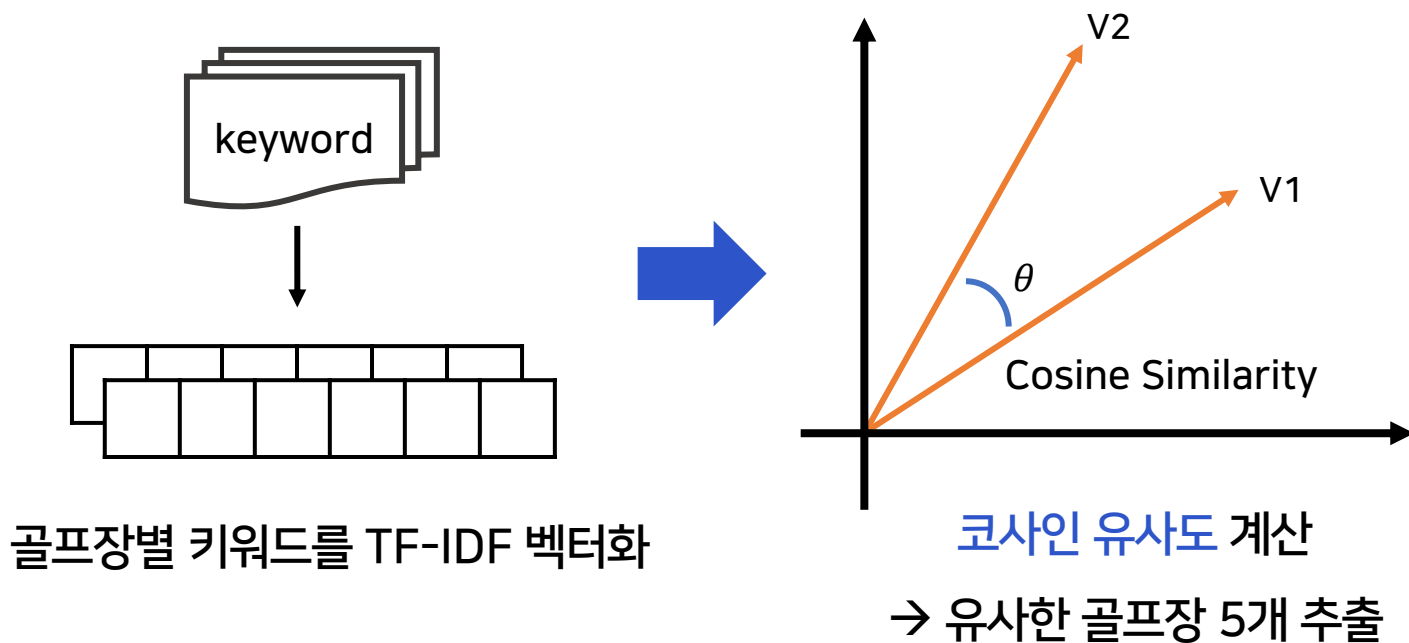
크롤링 데이터 전처리



회원권 유사도 계산



회원권 추천



01. 프로세스 - 회원권 추천



회원권 기본 정보



추가 데이터 수집



크롤링 데이터 전처리



회원권 유사도 계산



회원권 추천

기존 골프장을 입력하세요 : 레이테 .....> **골프장 입력**

기존 골프장을 입력하세요 : 레이테

=== 골프장 추천 순위 ===

	국가	골프장이름	일일체류비	그린피	카트비	캐디피	숙소	조식	중식	석식	식사	캐디팁	보험비	라커룸	송영	dist
0	베트남	빈펄	540000	1	1	1	1	1	0	0	0	1	0	0	1	2991.006627
1	베트남	옌중	417000	1	1	1	0	0	0	0	1	1	0	0	0	2656.037250
2	말레이시아	닐라이 스프링스	276000	1	1	0	1	1	1	1	3	0	0	0	0	4575.061433
3	베트남	치링	399200	1	1	1	1	1	0	0	1	1	0	0	0	2651.114670
4	중국	미션힐스	540000	1	1	1	1	1	0	0	1	0	1	1	0	2019.247427

=== 기준 골프장 : 레이테 ===

=== 추천된 골프장의 회원권 정보 ===

회원권번호	국가	골프장이름	상품명	기간	입회금	이용횟수	회원적용	비고
0	MN-01	말레이시아	닐라이 스프링스	VIP	평생	10000000	27.0	기명2 부부회원

키워드 기반 유사한 골프장 TOP 5

골프장의 거리와 일일 체류비를 고려한  
최종 추천된 골프장의 회원권 정보



## 02. 한계점

---

### 불충분한 데이터로 인한 한계점 존재

- 현재 보유하고 있는 회원권에 대한 기본적인 정보만으로 고객 맞춤형 서비스를 제공하기에는 데이터 부족
- 고객의 후기 및 평점 정보가 존재하지 않아 골프장 및 고객에 대한 세부적인 특성 고려 불가능
- 현재 회원권을 보유하고 있지 않은 고객, 즉 새로운 고객에 대해서는 추천 불가능

# 04

## 국내 골프장 추천 시스템

01. 프로세스

02. 시스템 시연

01. 프로세스 - Xgolf 후기 사이트 크롤링



데이터 수집



데이터 준비



모델링

01

순서	골프장	제목	개인만족도	평점/총평점	등록일자	조회/추천
237564	써닝포인트	그린,페어웨이 모두좋아요	만족	10.0 / 8.9	2021.08.24	2 / 0
237563	푸른솔강성	서비스좋은곳	보통	8.5 / 9.4	2021.08.24	12 / 0
237562	360도	비가와서	보통	7.5 / 8.7	2021.08.24	14 / 0
237561	리앤리	관리잘되었음	만족	10.0 / 8.9	2021.08.24	15 / 0
237560	파인힐스	후기	만족	10.0 / 9.6	2021.08.24	10 / 0
237559	파인힐스	가성비 최고였던 라운딩	만족	10.0 / 9.6	2021.08.24	11 / 0
237558	여주신라	골프	만족	10.0 / 9.4	2021.08.24	14 / 0
237557	더크로스비 야간	매우만족합니다	만족	10.0 / 9.1	2021.08.24	13 / 0
237556	동원썬밸리	동원썬밸리 좋아요	만족	10.0 / 8.3	2021.08.24	13 / 0
237555	알프스대영	알프스대영	보통	8.5 / 7.3	2021.08.24	17 / 0

Xgolf 후기

(<https://www.xgolf.com/board/afternote/list.asp>)

리뷰 수가 50개 초과 & 전체 평균 평점 9점 이상인  
골프장에 대해서만 추출

02

	골프장명	그린피(18홀 기준)	캐디피(4인 기준)	카트비(4인 기준)	식사(조식)	숙박(4인 기준)
0	현대더링스	240,000	130,000	90,000	NaN	NaN
1	오너스	390,000	150,000	100,000	NaN	NaN
2	샤인빌	170,000	130,000	100,000	NaN	NaN
3	타미우스	220,000	130,000	100,000	NaN	480,000
4	대명(비발디파크)	250,000	130,000	100,000	NaN	NaN
...	...	...	...	...	...	...
65	푸른솔포천	190,000	130,000	90,000	NaN	NaN
66	한림광릉	90,000	130,000	90,000	12,000	NaN
67	파인리즈	220,000	130,000	90,000	23,000	280,000
68	크리스탈밸리	105,000	130,000	100,000	20,000	200,000
69	클럽모우	190,000	130,000	100,000	NaN	NaN

추출한 골프장에 대한  
골프장별 일일체류비

01. 프로세스 - 추가 데이터 수집



데이터 수집

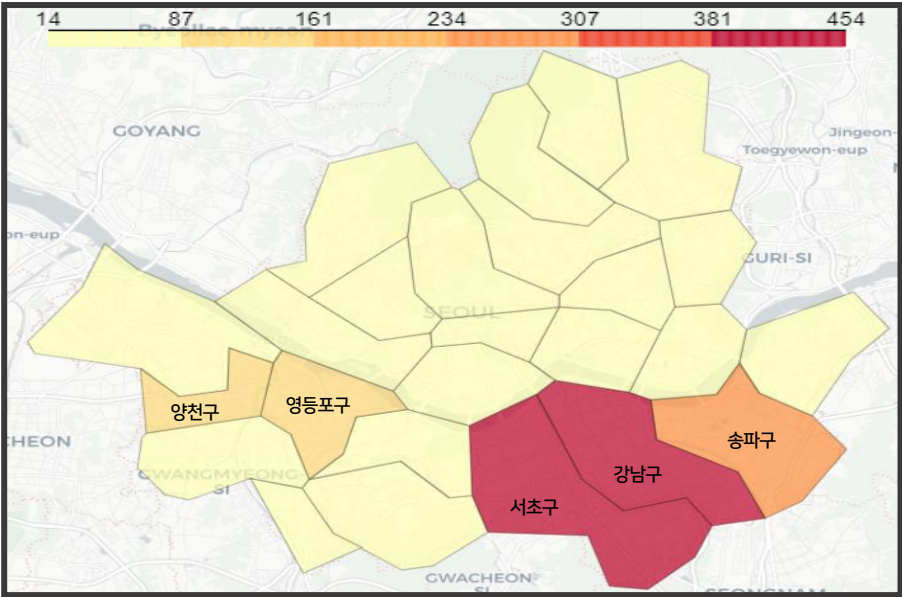


데이터 준비



모델링

03

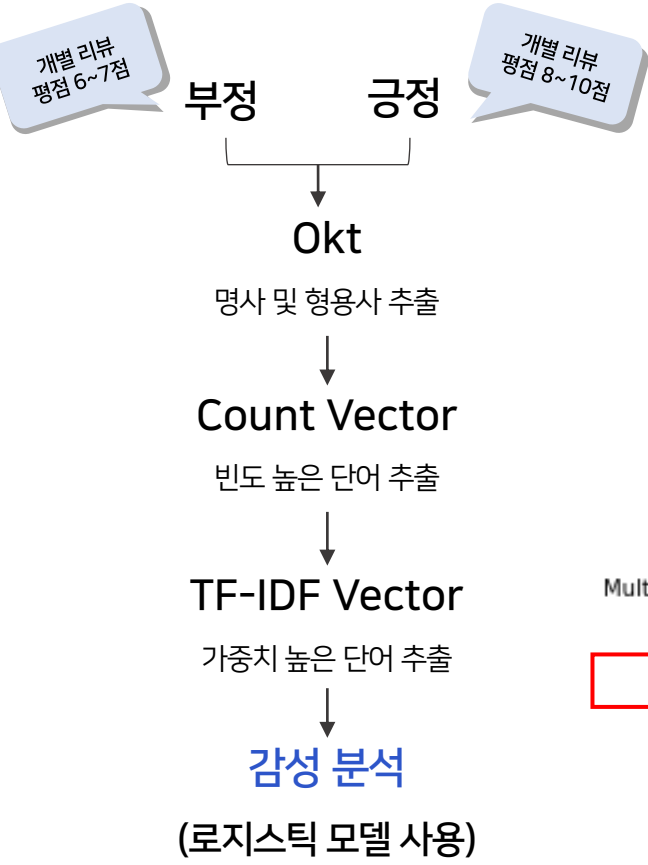


골프장명	위도	경도	distance	거리
현대더링스	36.72150968	126.3384081	(36.72150968, 126.3	105.7900563
오너스	34.83647102	128.0738565	(34.83647102, 128.0	310.4611779
샤인빌	33.34555148	126.7881264	(33.34555148, 126.7	462.2268334
타미우스	33.35087081	126.3867097	(33.35087081, 126.3	464.7674338
대명(비발디파크)	37.64139241	127.6752333	(37.64139241, 127.6	59.23300859
그린필드(제피로스)	33.46503779	126.6439973	(33.46503779, 126.6	449.774381
스프링데일	33.32952254	126.6380096	(33.32952254, 126.6	464.842416
360도	37.30274992	127.7388698	(37.30274992, 127.7	66.43507302
88	37.30508574	127.1624797	(37.30508574, 127.1	24.50612974
강남300	37.38565385	127.178311	(37.38565385, 127.1	18.21343247
골드	37.21770429	127.1369621	(37.21770429, 127.1	32.60594831

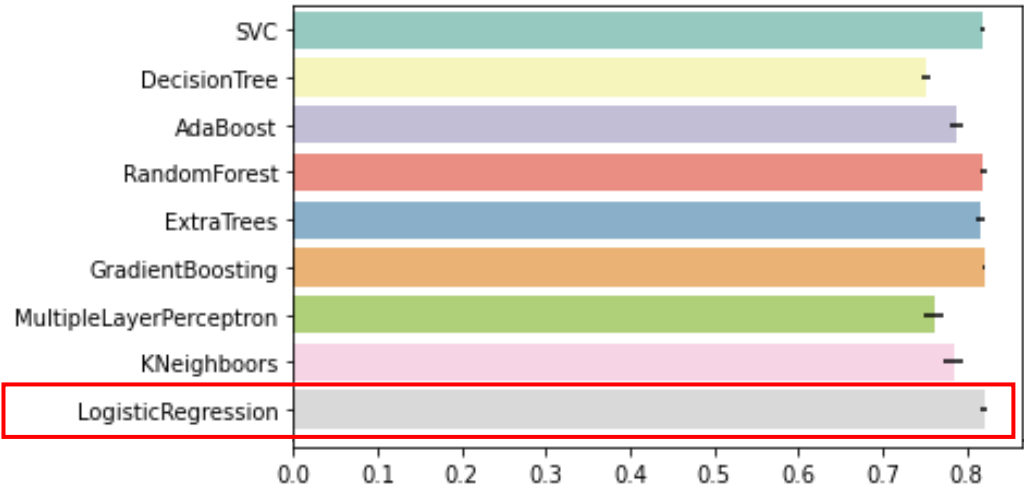
- 한케이골프의 회원은 강남구에 가장 많이 거주  
(주소 정보가 존재하는 5,873명 중 454명)
- 강남역과 각 골프장 간의 Haversine Formula를 활용한 최단 거리 계산

강남역 기준 골프장까지의 거리 계산

01. 프로세스 - 데이터 준비(감성분석)

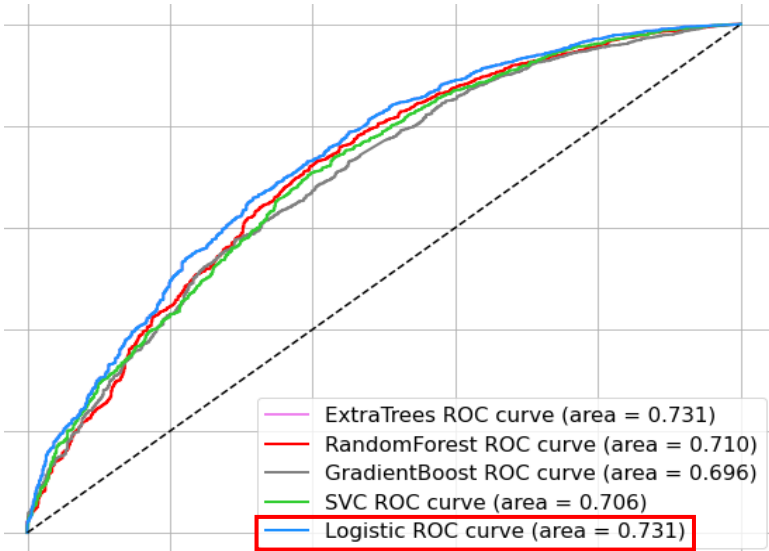


알고리즘별 교차검증의 평균 정확도



로지스틱 평균 정확도 : 0.83

ROC 곡선



로지스틱 ROC : 0.731

## 01. 프로세스 - 데이터 준비(Word2Vec)



데이터 수집



데이터 준비



모델링

	Link	Golf	Title	Date	Post	Caddie_Rate	Course_Rate	Price_Rate	ETC_Rate	Total_Rate
0	<a href="https://www.xgolf.com/board/afternote/view.asp...">https://www.xgolf.com/board/afternote/view.asp...</a>	레이캐슬	좋은 코스입니다.	2021.08.11	력서리한 골프장 아주 좋았어요. 티 간격도 10분으로 여유롭구요. 근처 양계장 냄새...	10	10	10	10	10.0
1	<a href="https://www.xgolf.com/board/afternote/view.asp...">https://www.xgolf.com/board/afternote/view.asp...</a>	임페리얼레이크	3부 노캐디	2021.08.11	3부예약이라 2인플레이 처음해봅니다^^ 자주 가는 임페리얼레이크 cc 라서 노캐디도 ...	10	10	10	10	10.0
2	<a href="https://www.xgolf.com/board/afternote/view.asp...">https://www.xgolf.com/board/afternote/view.asp...</a>	옥스필드	골프장 관리 곳 입니다.	2021.08.11	싼 가격에 부킹을 했기에 기대하지 않고 갔는데 념 관리가 잘되어 있어 기분 좋...	10	10	10	10	10.0



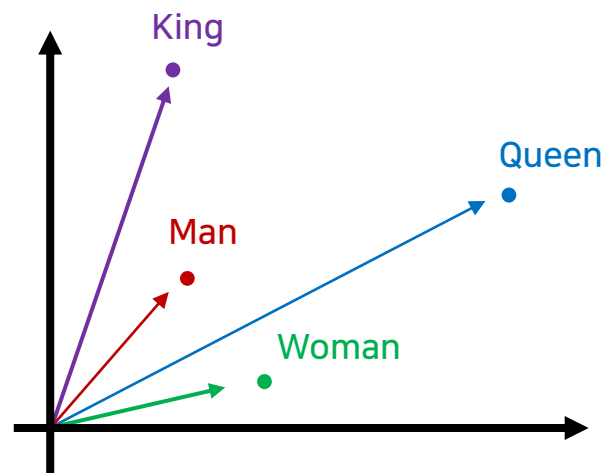
Okt

특수문자, 외국어, 조사, 한 글자 제외



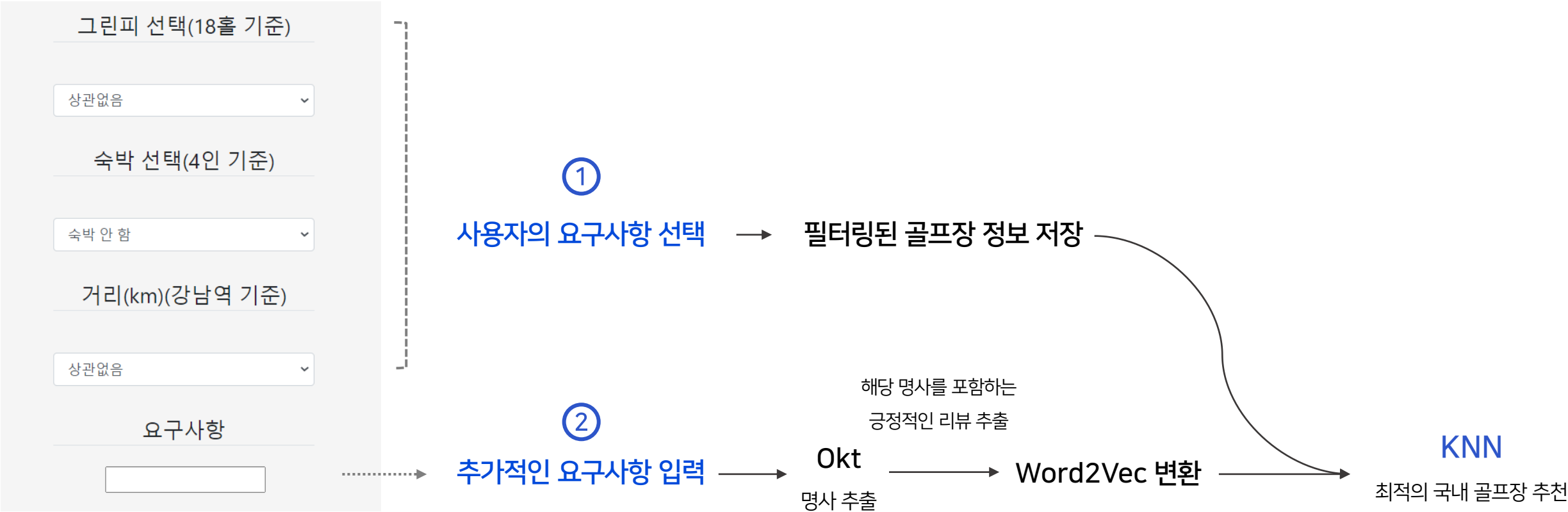
Word2Vec 학습

단어의 의미 벡터화

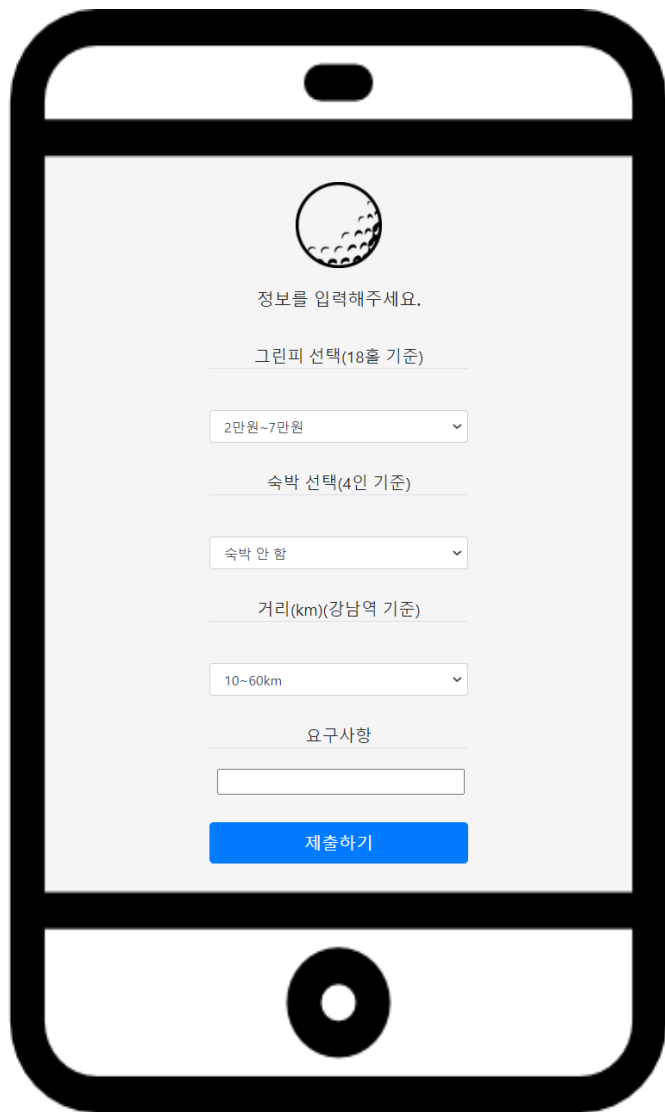


단어 간 유사도 반영

# 01. 프로세스 - 모델링



## 02. 프로그램 시연 - Flask를 활용한 추천 시스템 웹사이트 구현



A smartphone mockup displaying a golf course recommendation form. The form includes a circular profile picture placeholder, a text input field for "정보를 입력해주세요.", a "그린피 선택(18홀 기준)" label, a price range dropdown menu showing "2만원~7만원", a "숙박 선택(4인 기준)" label, a "숙박 안 함" dropdown menu, a "거리(km)(강남역 기준)" label, a distance range dropdown menu showing "10~60km", a "요구사항" label, and a text input field. A blue button labeled "제출하기" is at the bottom.

나에게 추천된 골프장은





# 05

## 분석 결과

- 01. 활용방안 및 기대효과
- 02. 새로운 영업 패러다임 제안

## 01. 서비스 활용방안 및 기대효과

회원권 추천 시스템

국내 골프장 추천 시스템

보유한 회원권의 특성과 유사한  
골프장의 회원권 추천 가능

고객의 선호도 반영 가능



고객이 관심을 가질 만한 회원권 추천 시스템 활용

## 02. 새로운 영업 패러다임 제안

### 기존 한케이골프의 영업 방식

- 전화 영업
- 가판 영업
- 신문 및 TV 광고
- 네이버 밴드를 활용한 홍보

고객의 선택권 고려 불가능



### 새로운 영업 패러다임

- 접근성이 좋은 온라인 플랫폼 활용
- 골프 커뮤니티에 추천 시스템 홍보

→ 한케이골프 홍보 및 유입 가능성 증대

시의적절한 고객 맞춤 정보 제공

# Q&A

# 감사합니다.

추천 시스템을 활용한 한케이골프 마케팅 제안

- 경기대학교 홀인원 팀 -

김재희 손현우 신다은 이세호 최세은 홍연정

