

A department recommendation service

---

# 진료과 추천 서비스

박장훈 송연정 유재현 정의석 양수빈





## 목차

---

01 소개

02 WBS

03 설계 구조

04 데이터 모델링

05 과정상 이슈

06 추후 개선 사항

## 서비스 소개

사용자가 건강에 관련된 문제를 겪을 때 어떤 진료과를 방문해야 하는지에  
대한 결정을 도와주는 서비스



## 서비스 기능

### - 증상으로 검색하는 기능

사용자가 증상을 입력하면 증상과 관련된 질병과 진료과를 추천하는 기능

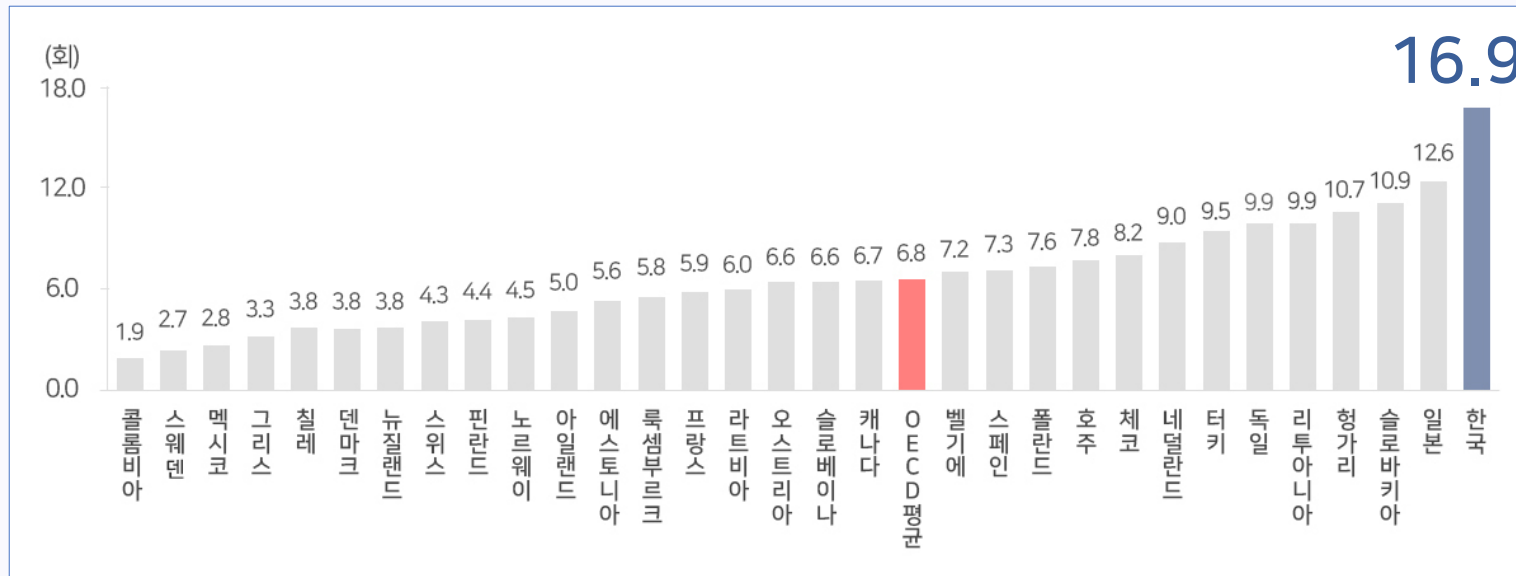
### - 이미지로 검색하는 기능

이미지에서 신체부위와 관련된 질병과 진료과를 추천하는 기능

# 기획 의도

## 1. 병원의 수는 많지만 진료과를 찾기 어려운 현실

### ■ OECD 국가 중 병원이용률이 가장 높은 대한민국



- 2022년 기준 연간 1인당 병원 방문 횟수 16.9회 → OECD 국가 중 가장 높은 수준

전국 전문과목별 병원·의원 통계(2024.4)

과목명	병의원 수
가정의학과	4,218
내과	7,779
마취통증의학과	2,848
비뇨의학과	1,787
산부인과	2,629
성형외과	1,417
소아청소년과	3,275
신경과	1,097
신경외과	1,581
안과	1,947
영상의학과	1,612
외과	3,107
이비인후과	3,058
재활의학과	1,598
정신건강의학과	2,368
정형외과	3,988
피부과	1,648

# 기획 의도

## 1. 병원의 수는 많지만 진료과를 찾기 어려운 현실

### 아프면 병원 가라는데 어느 병원 가야 돼

기사승인 2020-12-13 01:00:03

의외로 사람들이 잘 모르는 '증상별' 가야 할 병원... 싹 정리했습니다

2024-01-25 "통증환자 절반 엉뚱한 진료과 방문...개선 시급"

[메디칼타임즈=이인복 기자] 상당수 환자들이 만성 통증으로 고통받으면서도 통증클리닉이 아닌 타 병원이나 진료과를 찾아 질환을 악화시키고 있어 인식 개선이 시급하다는 지적이다.

전체 지식iN (1-10/2,333,165)

어느 병원



✓ 정확도 ✓ 최신순 ✓ 추천순

[스트레스성 턱관절은 어느 병원 가나요?... 한의사답변 2024.04.17](#)

... 아파서 병원 치료받아보려고 하는데요 한의원이나 한방병원에서도 치료를 받으시려는군요....

#스트레스 #턱관절 #턱관절통증

Q&A > 한방재활의학과 | 답변수 1 | 답변 dond\*\*\*\*

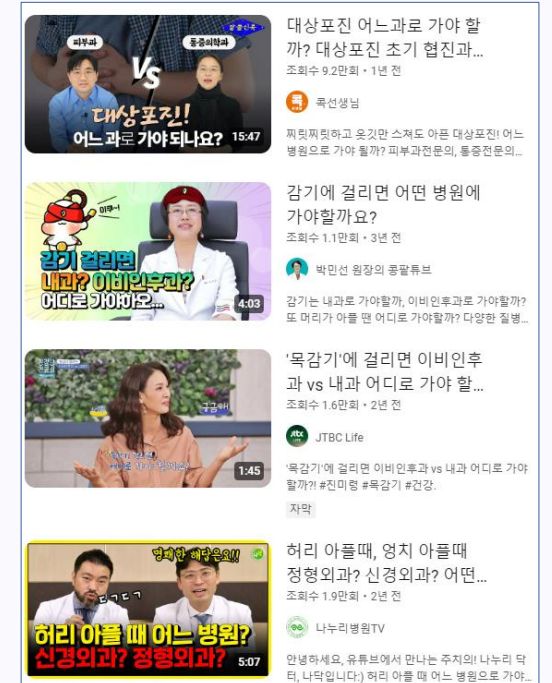
[턱장애 증상이라면 어느 병원에 가야?... 한의사답변 2024.01.24.](#)

적절한 병원을 찾으려는 수요는 꾸준히 증가

어느 병원, 어떤 병원, 어디 병원, 무슨 병원 등

➡ 병원 진료과를 찾는 질문건수 누적 평균 2~300만 건

➡ 일 평균 약 3,000건의 질문 등록



# 기획 의도

## 2. 접근성이 좋은 병원을 선호하는 경향

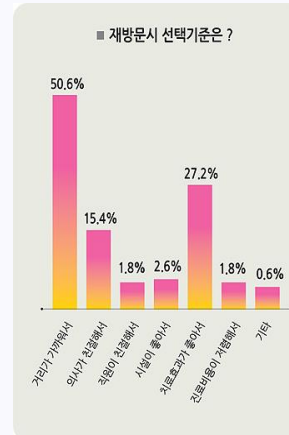
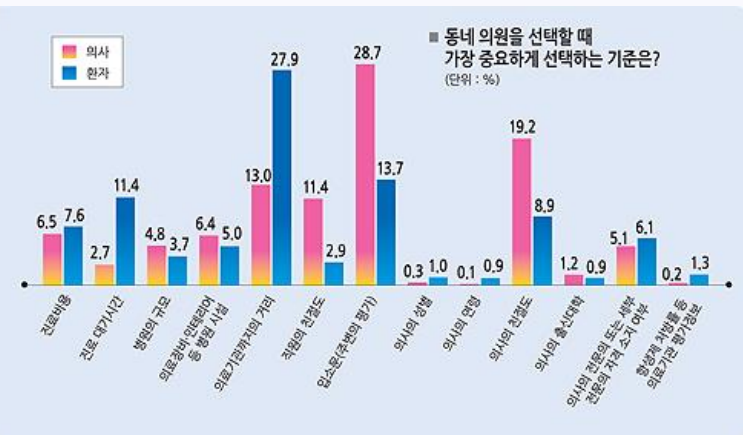
### 병원 선택의 이유, 1위는 '지리상 이점'

청년의사 \* 이민주 기자 | ☎ 입력 2018.10.04 12:00 | 💬 댓글 0

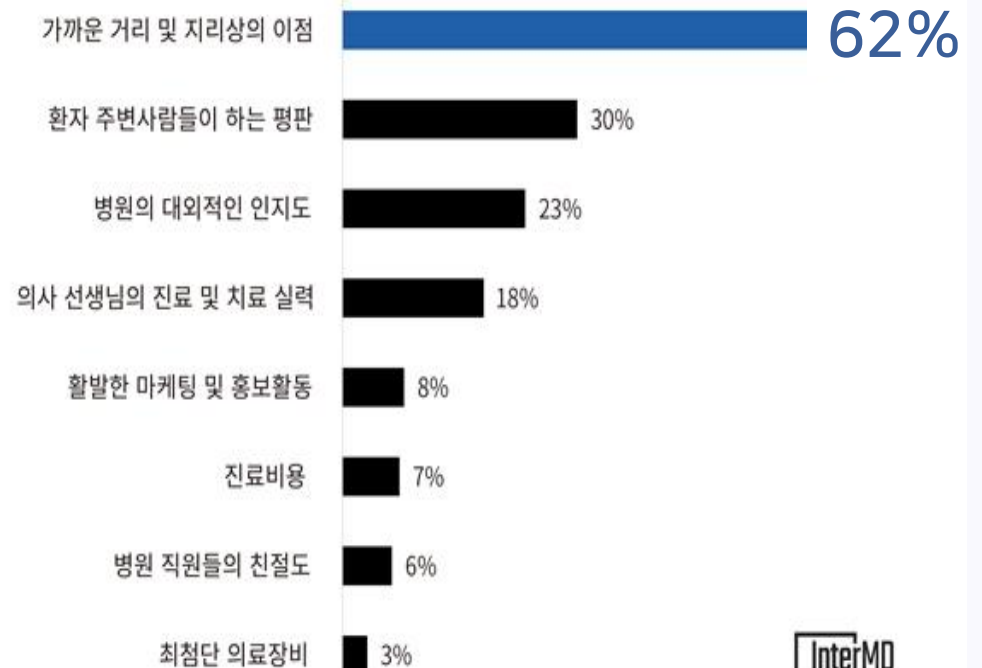
국민은 의원급 의료기관을 이용할 때 '가까이에 있는 의료기관' 즉, '접근성'을 가장 중요하게 생각하는 선택기준으로 꼽은 것이다. 이같은 응답은 500명이 응답한 1순위 조사에서 더 두드러진다. '의료기관까지의 거리'가 42.8%(214명)로 다른 항목보다 월등히 높았다.

| 62%가 가까운 거리가 영향준다고 응답...평판 30%, 실력 18%

의사와 환자 모두가 병원을 선택하는데 있어 '지리적 이점'을 가장 큰 요인으로 고려한다는 조사 결과가 나왔다.



### 환자가 병원을 선택하는 요인



InterMD

(전체: 인터엠디 의사회원 501명, 단위: 중복 %)

# 기획 의도

## 2. 접근성이 좋은 병원을 선호하는 경향

### 병원선택 요인이 고객만족과 재이용의도에 미치는 영향

The Effects of Hospital Choice Factors on Customer Satisfaction and Revisiting Intention in General Hospitals

미치는 영향을 분석하였다. 본 연구의 결과를 선행연구와 비교해 보면, 병원선택요인 중 인적요인, 접근성요인, 대외이미지요인이 고객만족에 유의한 영향을 미치는 것으로 검증되었다. 이는 병원선택요인 및 만족도 제고 연구에서 인적전문성, 접근편리성, 신속간편성이 환자만족에 큰 영향을 미친다는 이경애[14]의 연구, 병

### 초진과 재진환자의 종합병원 선택요인 비교분석

양종현  
중원대학교 보건행정학과

### A Comparative Study on Hospital Choice Factors between First-visit and Revisit patients in General Hospitals

둘째, 초진과 재진환자의 병원선택요인 비교분석에서 초진환자의 경우 모형1과 모형2에서 인적요인, 효율적요인, 접근성요인이 유의한 양(+)의 영향력을 미쳤다. 재진환자의 경우 모형1과 모형2에서 인적요인, 물리적요인, 접근성요인, 대외이미지요인이 유의한 양(+)의 영향력을 미쳤다. 그리고 모형 2에서는 소득이 유의한 양(+)의 영향력으로 나타났다. 이는 초진, 재진환자의 고객접점

소개

WBS

설계 구조

데이터 모델링

과정상 이슈

추후 개선 사항

# WBS

작업 / 기간	4월 1주차	2주차	3주차	4주차	5월 1주차	2주차	3주차
서비스 기획	→						
데이터 수집		→					
데이터 모델링		→	→				
서비스 구현				→	→	→	
테스트						→	
문서 작성							→

- 서비스 기획 : 주제 선정, 설계 구조 계획, 요구사항 정의 etc.
- 데이터 수집 : 질환/증상 데이터 수집, 네이버 지식인 데이터 크롤링, 공공데이터 수집
- 데이터 모델링 : 자연어 처리와 추천시스템 모델링
- 서비스 구현 : DB 구축, 기능 구현 및 웹 사이트 연동
- 테스트 : 서비스 QA 및 기타 기능 구현
- 문서 작성 : PT 자료 준비



# 개발 환경

협업



Front  
Back

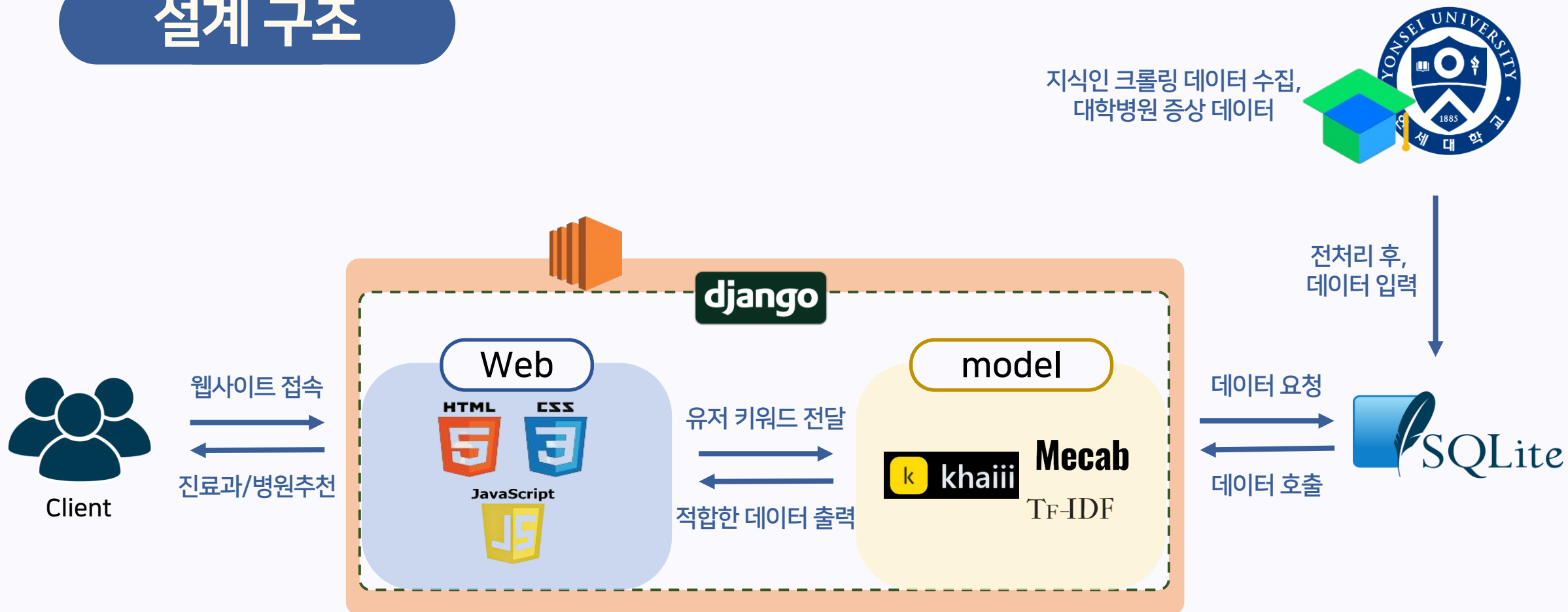


개발환경



DB

# 설계 구조



# ERD

Disease	
PK	<u>병명 id</u>
	병명
	증상
	발병 부위
	진료과
	관련 태그

- Id, 병명, 증상, 발병 부위, 진료과, 관련 태그 총 6가지의 컬럼으로 구성된 테이블 하나로 DB를 구성
- 초기 기획 시 4 개의 테이블로 구상 했으나, 병원 추천 기능과 관련하여 DB를 구성하는 것보다 길 찾기 API를 활용하는 것이 유용할 것이라 판단되어 기획 수정

\*Disease table

병명 id	발병 부위	증상	병명	진료과	관련 태그
1	복부	신생아의 98%는 출생 후 24시간 내에 짙은 녹색의 태변을 배설합니다. 아기의 항문이 정상적으로 발달되어 있음에도 24시간이 지나도록 태변을 간에서 만들어진 담즙은 담도를 통해 소장으로 분비돼 소화를 돕는 데 사용됩니다. 그런데 담도가 막히면 담즙이 분비되지 못하고 간에 정체되면	선천성 거대결장	소아외과	히르쉬스프루그병, 선천성대장질환, 거대결장, 대장조영술검사, 담도, 담도폐쇄증, biliary atresia
2	복부		담도폐쇄증	소화기내과	간경변증, 간이식, 담도조영술, 피로감, 구토, 복통, 황달, 메스꺼움, 식욕부진, 체중감소
3	복부	조기 간암의 증상으로 피로감, 전신 쇠약감, 소화 불량, 상복부에 느껴지는	간암	간담체외과	복통, 설사, 식욕부진, 체중감소
4	복부	대표적인 증상으로는 복통, 설사, 체중 감소를 들 수 있습니다. 특히 10~20대 젊은 연령대에서 1~2개월 이상 배 아픔과 설사가 지속되면 크론	크론병	소화기내과	소, 치루, 장폐색, 국소성 장염, 비만, 고도비만, 비만대사수술, 루와이위우회술, 위소매절제술,
5	복부	비만은 체내 근육량에 비해 지방조직이 과다한 상태를 의미합니다. 대한비	고도비만	내분비내과	

기능 설명



### 증상으로 검색

현재 겪고 계신 증상을 입력하시면, 해당 증상과 관련한 질병들을 알아볼 수 있어요.

증상을 적어주세요.

검색 하기



### 이미지에서 선택하기

이미지 상에서 특정 신체 부위를 선택하면 해당 부위와 관련된 질병을 알아볼 수 있어요. 겪고 계신 증상이 어떤 부위와 관련이 있는지 파악하고, 그에 따른 질병을 알고 싶을 때 유용하게 사용할 수 있어요.

소개

WBS

설계 구조

데이터 모델링 ●

과정상 이슈

추후 개선 사항

# 데이터 모델링

## 데이터 수집

- Selenium

## 데이터 전처리

- Mecab, Khaiii

## EDA

## 데이터 모델링 – Crawling Data



# 데이터 수집



세브란스병원  
SEVERANCE HOSPITAL

Data Info.	연세대 세브란스 병원 데이터
데이터 수	334
컬럼 수	5
컬럼 종류	발병 부위, 증상, 질환명, 진료과, TAG



## 데이터 수집



Data Info.	네이버 지식iN 데이터
데이터 수집	Selenium
데이터 수집 기간	2010.09 - 2024.04
데이터 크기	11.59 GB
데이터 수(row)	3,722,534
컬럼 수	4
컬럼 종류	제목, 질문, 답변, URL

**Q.** 가만히 앉아있다가 어느순간 눈 앞이 0.1초간 정전된 것 같은 걸 느껴요.

바람 · 조회수 1,121 · 2018.12.23

가만히 앉아있다가 어느순간 눈 앞이 0.1초간 정전된 것 같은 걸 느껴요. 병원에 가야할까요? 만일 가야한다면, 어느 병원으로 가야하나요?

신경과



s450\*\*\*\*님,  
정보를 공유해 주세요.

답변

1 개 답변

최적 ✓ 추천순 ✓



오경필

전문외

메디인병원

본인 입력 포함 정보 ⓘ

안녕하세요. 하이닥-네이버 지식iN 상담의 신경과 전문의 오경필 입니다.

말씀하신 증상만으로는 구체적인 답변이 어렵습니다.  
하지만, 말씀하신 증상이 시야 장애를 동반하거나, 찢어진 양상의 통증을 동반하는 경우에는 뇌혈관질환, 뇌실질질환 및 삼차신경통 등 여러 가능성에 대한 고려가 필요합니다.

이러한 질환을 담당하는 과는 신경과이므로,  
증상이 반복되거나 불편감이 지속된다면, 진료를 보시길 권해드립니다.

질문자님의 자세한 병력, 진찰 상태를 알수 없어 짧게 답변 드리는 점 양해바라며, 도움되시길 바랍니다.  
 쾌유 바랍니다. 감사합니다.

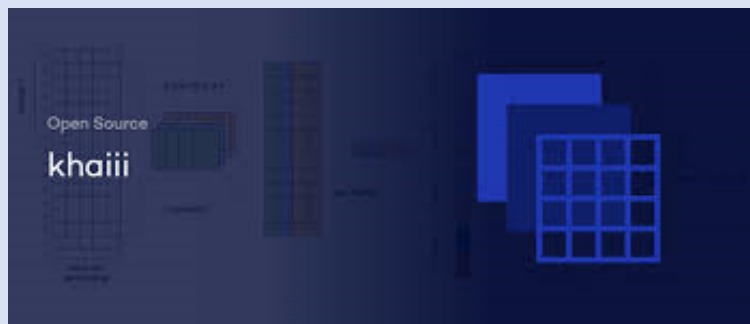
☰ 담당키워드	☀ 2024	☀ 2023
무슨 병원	● 완료	● 완료
무슨 진료과	● 완료	● 완료
병원 어디	● 완료	● 완료
어느 병원	● 완료	● 완료
어느 진료과	● 완료	● 완료
어디 병원	● 완료	● 완료
어디 진료과	● 완료	● 완료
어떤 병원	● 완료	● 완료
어떤 진료과	● 완료	● 완료
진료과 어디	● 완료	● 완료



# Preprocessing

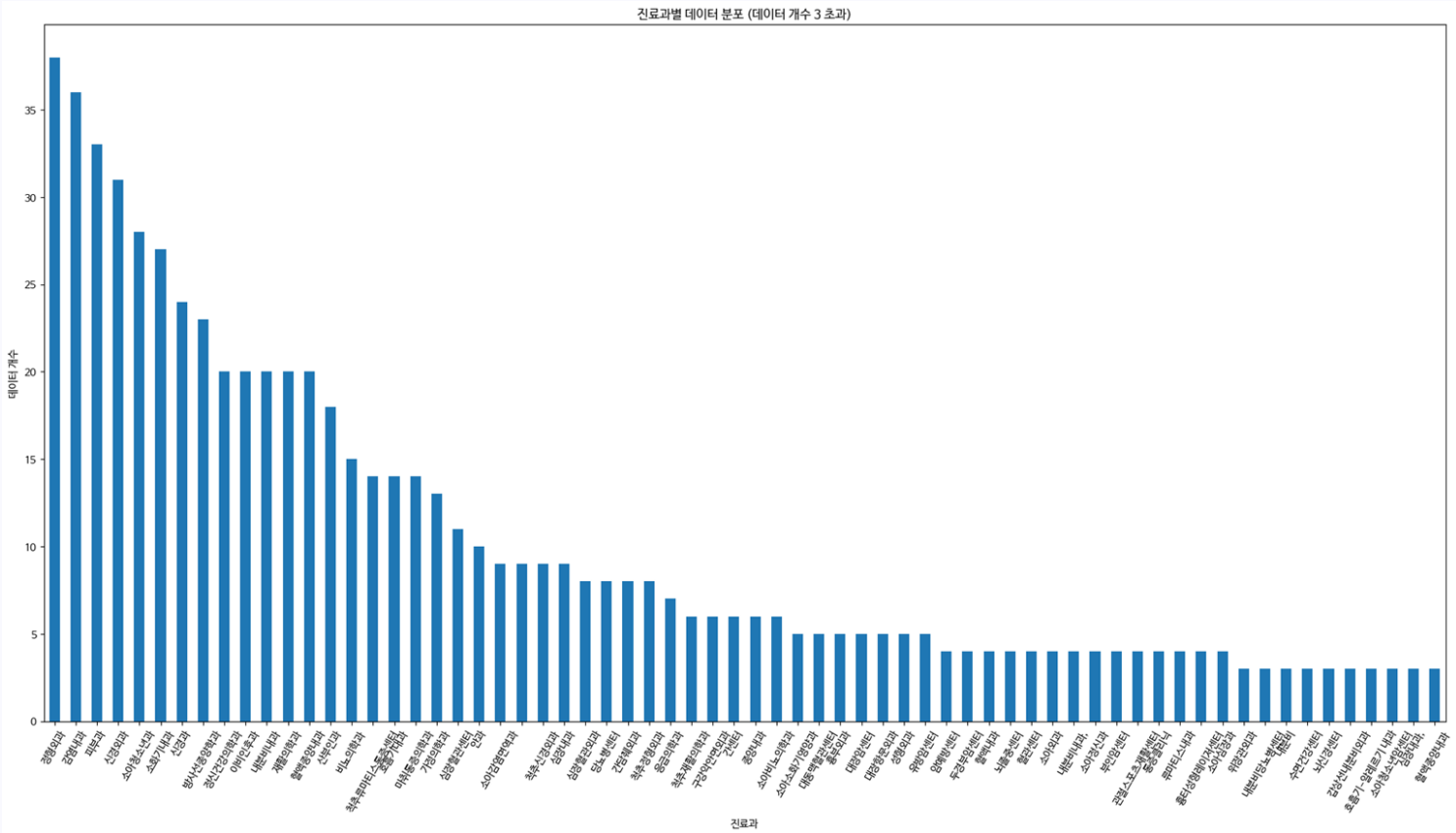
## Mecab

0 [신생아, 출생, 후, 시간, 내, 녹색, 태변, 배설, 아기, 항문, 정상, 발달...  
1 [간, 담즙, 담도, 소장, 분비, 소화, 데, 사용, 담도, 담즙, 분비, 간, ..  
2 [조기, 간암, 증상, 피로감, 전신, 쇠, 약감, 소화, 불량, 상복, 부, 체중...  
3 [대표, 증상, 복통, 체중, 감소, 수, 대, 연령, 개월, 이상, 배, 아픔, ...  
4 [비, 체내, 근육, 량, 지방, 조직, 과다, 상태, 의미, 비, 학회, 신체, ...  
...  
328 [두개골, 세, 이하, 소아, 수두증, 뇌척수액, 축적, 머리, 머리, 둘레, 정수...  
329 [시야, 빛, 아지랑이, 것, 시각, 통, 구증, 하품]  
330 [발열, 피로감, 몸살, 기운, 기침, 구토, 식욕, 감소, 통, 목, 의식, 장애...  
331 [주, 잠복기, 대부분, 무증상, 열, 통, 어지럼증, 구토, 의식, 혼미, 경련]  
332 [미열, 통, 피로감, 목, 뒤, 신경, 마비, 의식, 장애, 경련]



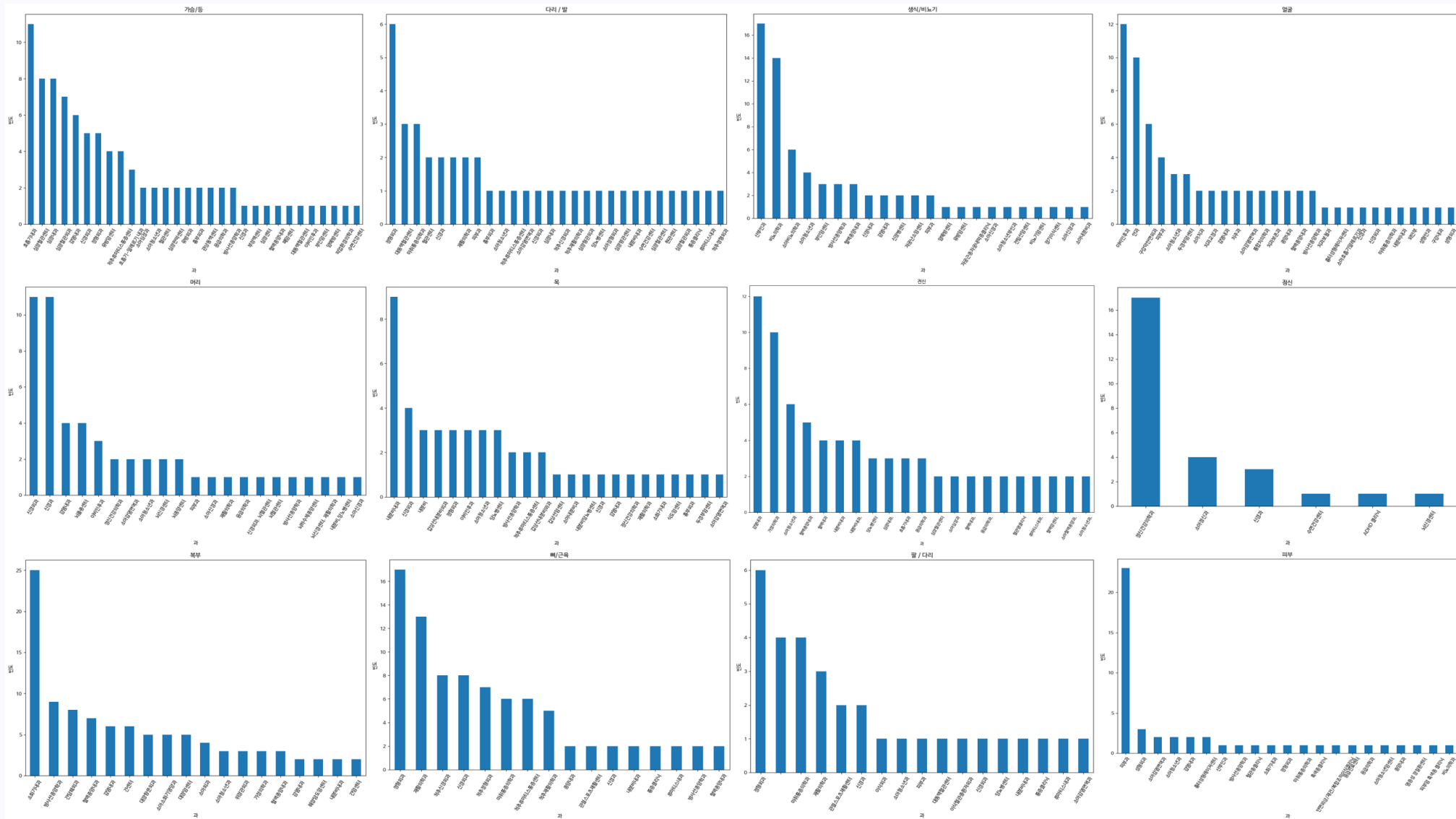
0 신생아 출생 시간 녹색 태변 배설 아기 항문 정상 발달 시간 태변 배설 선천 거대 ...  
1 담즙 담도 소장 분비 소화 사용 담도 담즙 분비 정체 손상 간경변증 진행 담도 폐쇄...  
2 조기 간암 증상 피로감 전신 쇠약감 소화 불량 상복부 체중 감소 증상 간암 진행 기...  
3 대표 증상 복통 체중 감소 연령 이상 아픔 설사 지속 크론병 가능 의심 체중 감소 ...  
4 비만 체내 근육량 지방 조직 과다 상태 의미 한비만학 신체 질량 지수 이상 비만 이...  
...  
328 개골 이하 소아 수두증 뇌척수액 축적 머리 머리둘레 정수리 대천문 기력 구토 발작 ...  
329 시야 아지랑이 시각 조짐 두통 전구 증상 하품  
330 발열 피로감 몸살 기침 구토 식욕 감소 두통 의식 장애 경련  
331 잠복기 대부분 증상 고열 두통 어지럼증 구토 의식 혼미 경련  
332 미열 두통 피로감 신경 마비 의식 장애 경련

## EDA-Severance data

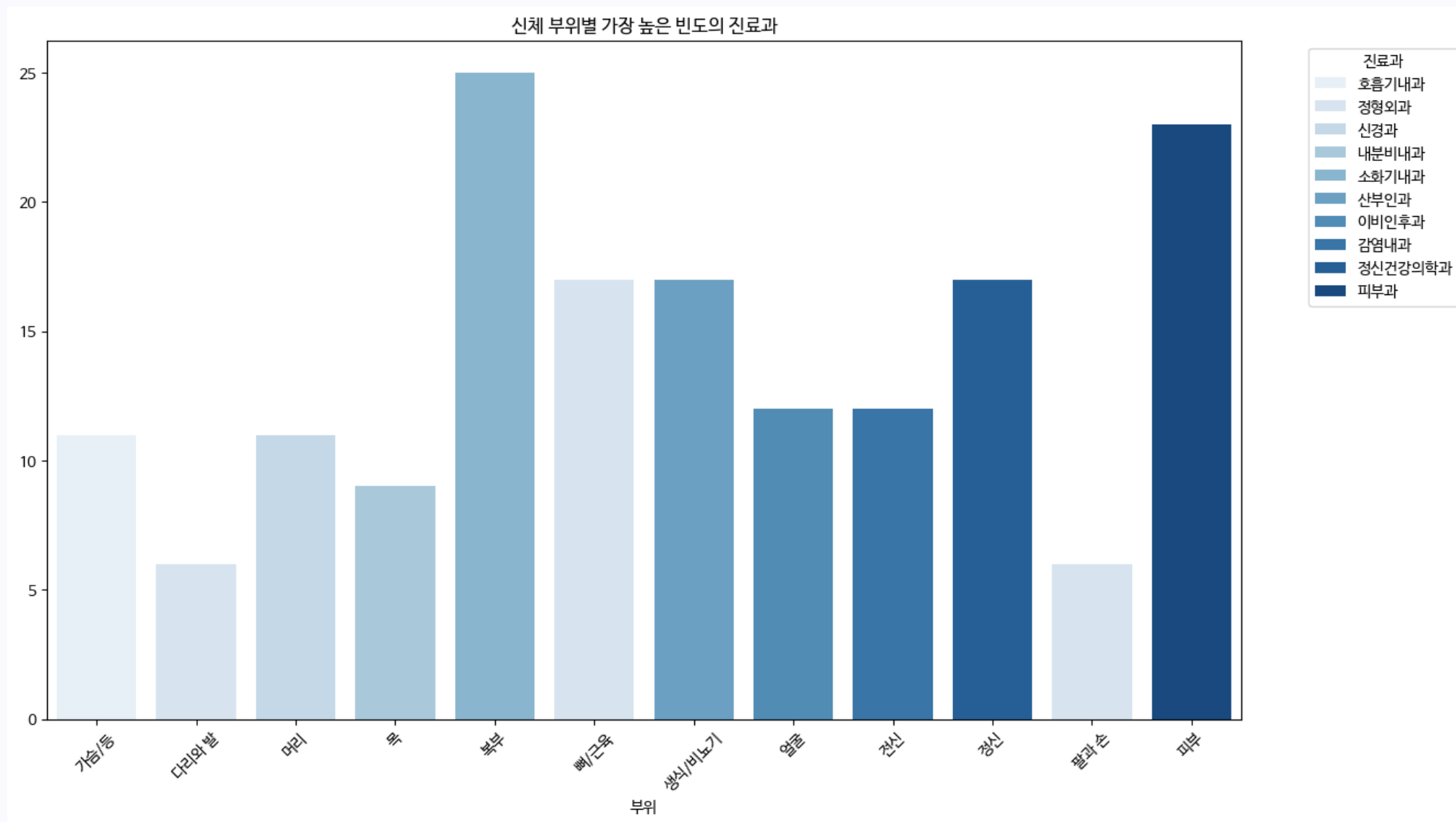


진료과	
정형외과	38
감염내과	36
피부과	33
신경외과	31
소아청소년과	28
	..
चेदददददददद	1
호흡기-알레르기내과	1
호흡기내과,	1
알레르기-천식센터,	1
심장혈관외과	1

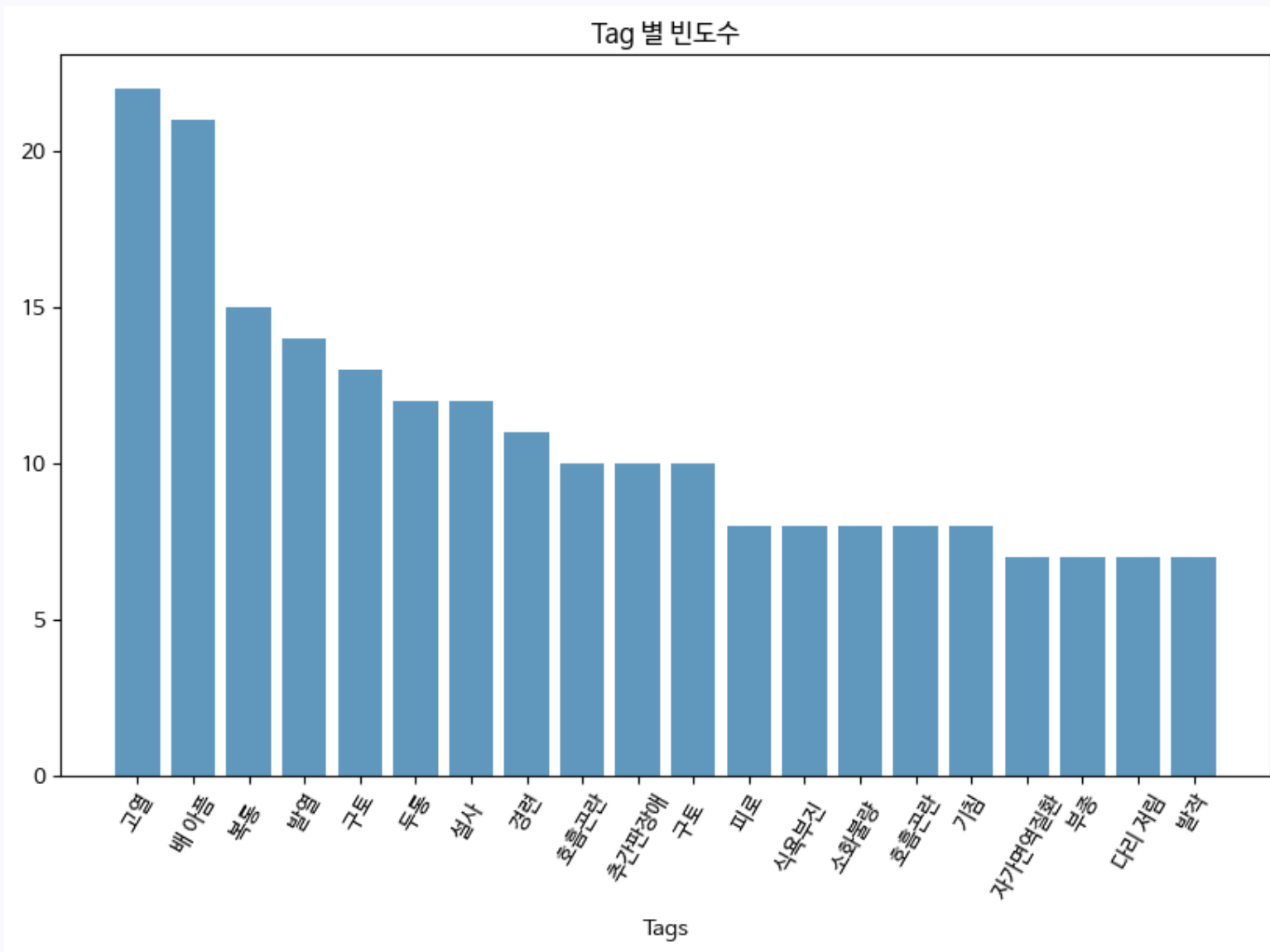
## EDA-Severance data



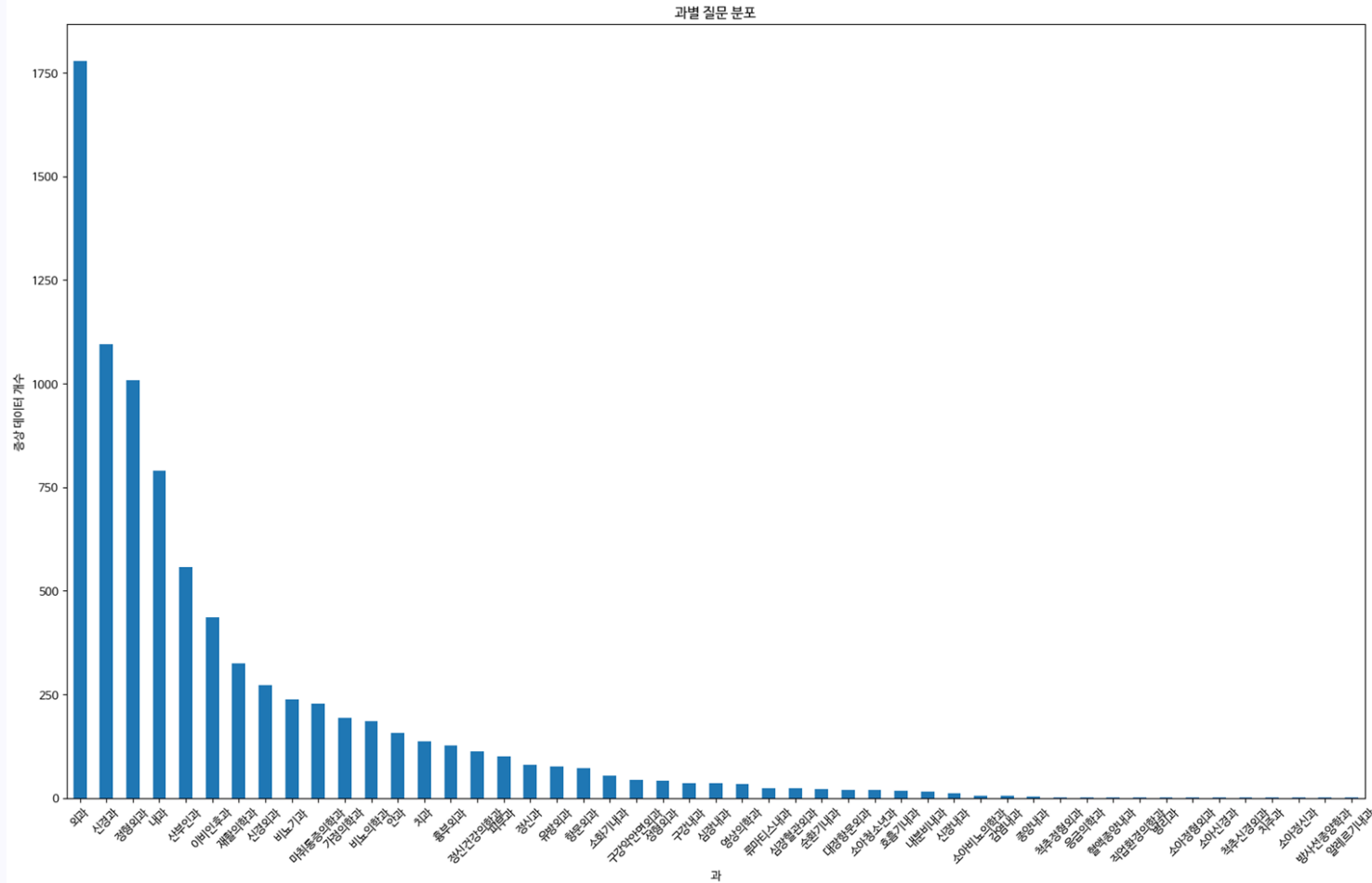
# EDA-Severance data



# EDA-Severance data



# EDA-Crawling data



외과	1779
신경과	1094
정형외과	1008
내과	790
산부인과	558
이비인후과	436
재활의학과	324
신경외과	273
비뇨기과	238
마취통증의학과	228
가정의학과	193
비뇨의학과	186
안과	156
치과	136
흉부외과	127
정신건강의학과	112
피부과	100
정신과	80
유방외과	75
항문외과	71
소화기내과	53
구강악안면외과	44
성형외과	41
구강내과	35
심장내과	35

영상의학과	33
류마티스내과	23
심장혈관외과	23
순환기내과	22
대장항문외과	20
소아청소년과	20
호흡기내과	18
내분비내과	15
신장내과	11
소아비뇨의학과	5
감염내과	5
종양내과	3
척추정형외과	2
응급의학과	2
혈액종양내과	2
직업환경의학과	2
병리과	1
소아정형외과	1
소아신경과	1
척추신경외과	1
치주과	1
소아정신과	1
방사선종양학과	1
알레르기내과	1

## 모델에 사용한 알고리즘

알고리즘	Accuracy
RandomForest	0.211
Support Vector Classifier	0.328
XGBoost	0.189
RNN	0.136
Cosine Similarity	-

$$\begin{aligned} \text{similarity} &= \cos \theta = \frac{A \cdot B}{\|A\| \|B\|} \\ &= \frac{\sum_{i=1}^n A_i \times B_i}{\sqrt{\sum_{i=1}^n (A_i)^2} \times \sqrt{\sum_{i=1}^n (B_i)^2}} \end{aligned}$$

증상을 입력하세요: 불면증이 있고 집중력이 떨어지는 증상이 자주 있습니다  
추천 과: ['가정의학과', '정신과']

유사도: 0.5029

증상을 입력하세요: 무릎 관절이 아픕니다.  
추천 과: ['정형외과', '외과']

유사도: 0.6798

소개

WBS

설계 구조

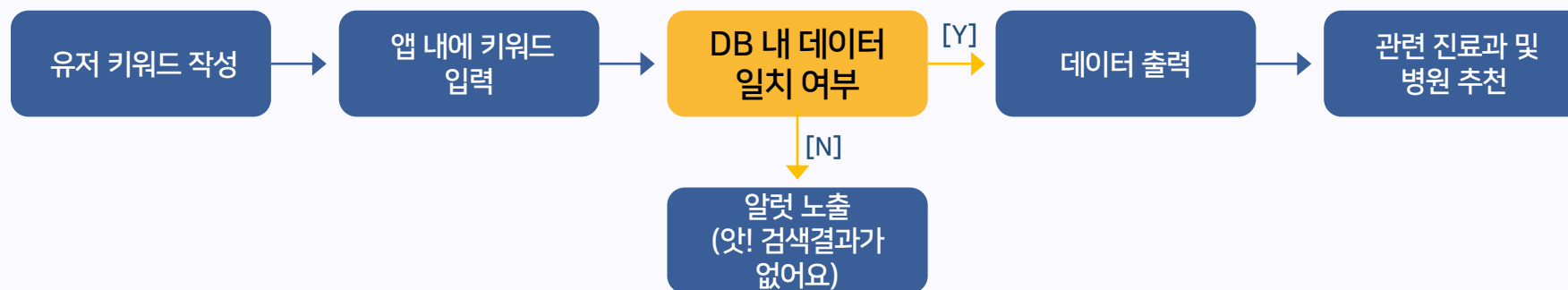
데이터 모델링

과정상 이슈

추후 개선 사항

## 과정상 이슈

- DB에서 사용자가 기입하는 데이터를 DB에서 꺼내오는 과정에서, 이를 인식하지 못하여 출력이 없는 문제가 발생



- 이를 해결하기 위해, TF-IDF가중치를 활용한 코사인 유사도 분석을 통해 해결함



# TF-IDF

※ 단어의 빈도와 역 문서 빈도를 사용하여 DTM 내의 각 단어들마다 중요한 정도를 가중치로 주는 방법

$$idf(t) = \log\left(\frac{n}{1 + df(t)}\right)$$

TF : 특정 문서에서 특정 단어의 등장 횟수

IDF : DF(특정 단어가 등장한 문서의 수)에  
반비례하는 수

- 모든 문서에서 자주 등장하는 단어는 중요도가 낮다고 판단하며, 특정 문서에서만 자주 등장하는 단어는 중요도가 높다고 판단.
- 해당 값이 작으면 중요도가 낮고, 값이 크면 중요도가 높은 것
- 따라서 불용어의 경우, 다른 단어에 비해 낮은 중요도를 갖게 된다.

# Cosine similarity

※ 두 벡터 간의 코사인 각도를 이용하여 구할 수 있는 두 벡터의 유사도를 의미

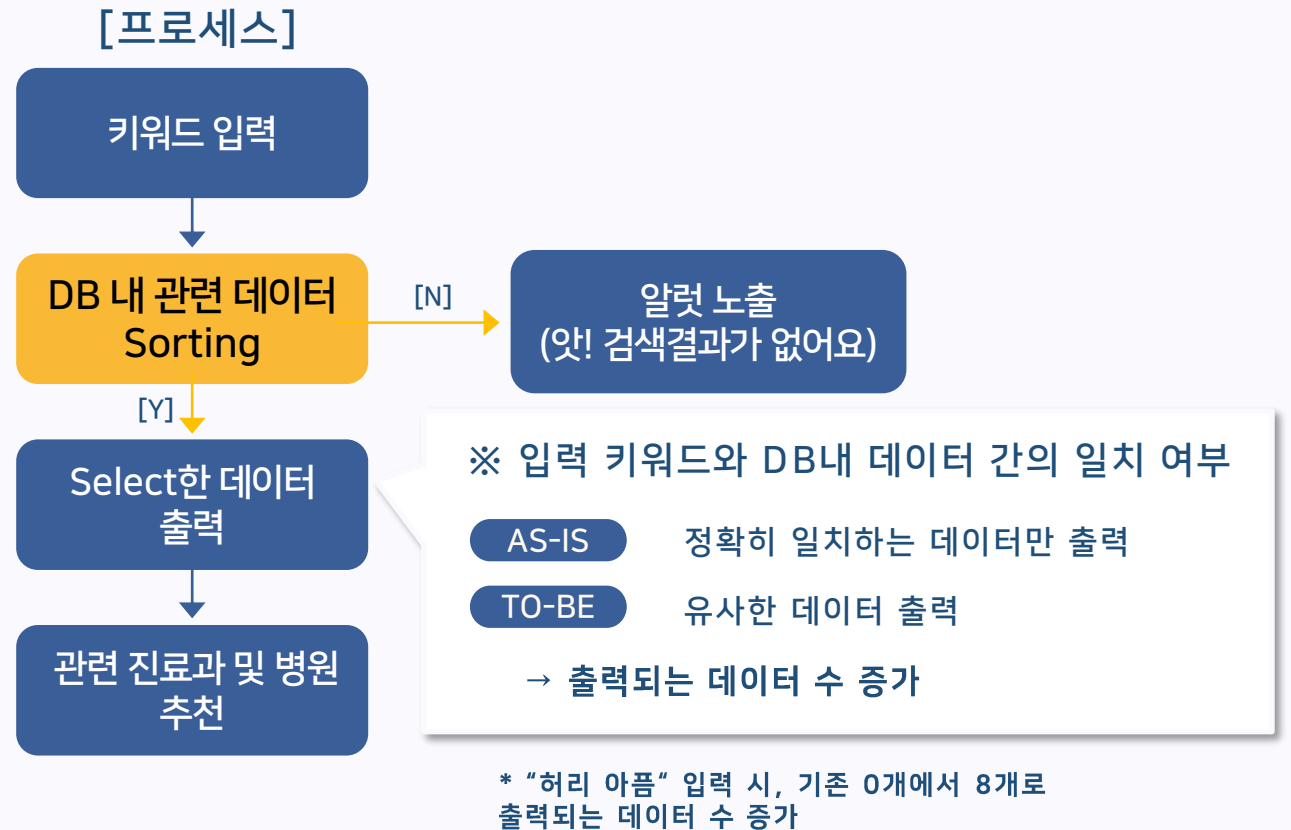
$$\begin{aligned} \text{similarity} = \cos \theta &= \frac{A \cdot B}{\|A\| \|B\|} \\ &= \frac{\sum_{i=1}^n A_i \times B_i}{\sqrt{\sum_{i=1}^n (A_i)^2} \times \sqrt{\sum_{i=1}^n (B_i)^2}} \end{aligned}$$

- 코사인 유사도는 -1이상 1이하의 값을 가지며, 값이 1에 가까울수록 유사도가 높다고 판단할 수 있다.
- 이를 직관적으로 두 벡터가 가리키는 방향이 얼마나 유사한가를 의미한다.

# 해결 과정

## 왜 이것을 사용하였는가?

- 보통 증상에 대한 설명은 확보한 데이터에서 1~2회 가량 설명하는데, TF-IDF를 사용했을 때 가중치가 높게 나온다.
- 그리고 해당 값을 활용해 코사인 유사도를 구한다면 입력 키워드와 유사한 내용이 들어간 데이터를 추출하기 수월해지기 때문에 사용



소개

WBS

설계 구조

데이터 모델링

과정상 이슈

추후 개선 사항 ●

## 새롭게 추가할 기능

### 로그 저장을 위한 EFK 활용

- 유저가 입력한 키워드, 출력데이터에 대한 만족도 파악, 통계 정보 등 앱의 전반적인 상태를 파악하기 위함

### 설문을 통한 진료과 추천

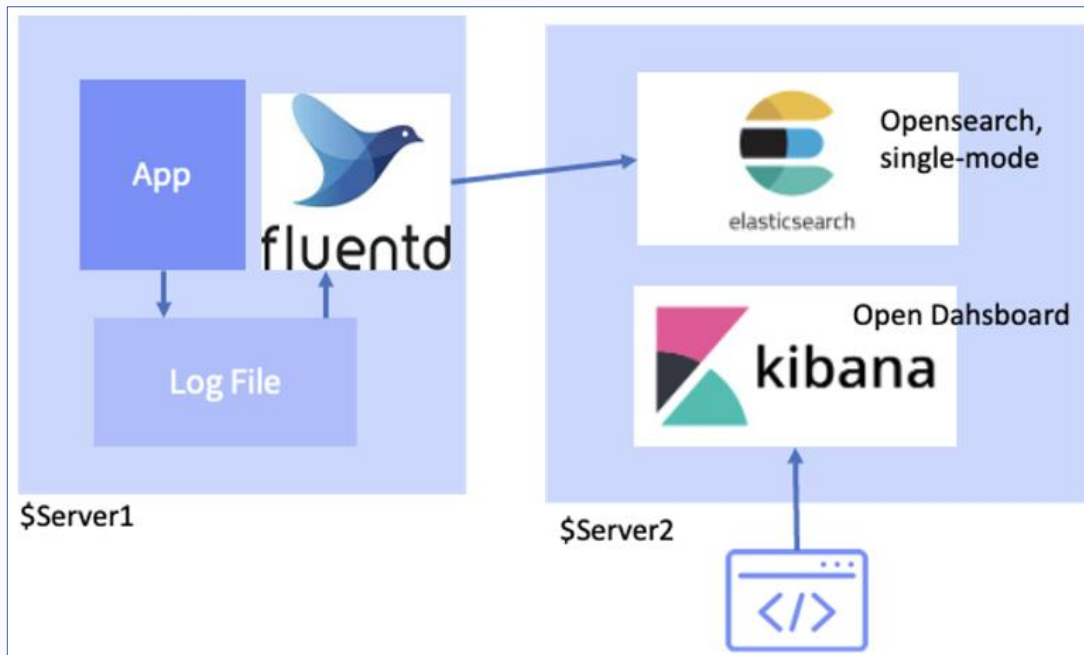
- 최초 설계 시, 기획했으나 구현하지 못했던 기능

## 개선할 요소

### 미구현 모델의 보완

- 크롤링 데이터를 바탕으로 만들어진 추천 모델 구현

# 로그 저장을 위한 EFK 활용



- 사용자가 입력한 키워드, 출력데이터에 대한 만족도 파악을 위해 로그를 수집하고자 위함
- 통계 정보 등 앱의 전반적인 상태를 파악하여 추후 개선 및 개발

# 설문을 통한 추천 기능

×

"관절에 문제가 있으신가요?"

☒ 해당 증상이 없어요

☐ 관절이 잘 움직이지 않아요

☐ 관절 모양이 이상해요

☐ 관절이 부었어요

☐ 관절염이 있어요

☐ 턱관절이 이상해요

☐ 관절이 아파요

선택완료

×

"증상이 언제 시작되었나요?"

11 12 13 14 15 16 17

일 전부터

선택완료

×

"질병을 진단하는 서비스가 아닌 증상 유사도를 분석하여 질병 정보를 제공하는 서비스이에요! 증상이 심하면 꼭 의료기관에 방문해주세요."

이런 사람은 이용할 수 없어요!

응급환자

임산부

코로나 19 의심환자

입력한 증상들을 AI가 학습하고 있어요

증상체크 시작하기

# 미구현 기능 보완



## 증상으로 검색

현재 겪고 계신 증상을 입력하시면, 해당 증상과 관련한 질병들을 알아볼 수 있어요.



## 이미지에서 선택하기

이미지 상에서 특정 신체 부위를 선택하면 해당 부위와 관련된 질병을 알아볼 수 있어요. 겪고 계신 증상이 어떤 부위와 관련이 있는지 파악하고, 그에 따른 질병을 알고 싶을 때 유용하게 사용할 수 있어요.

```
(django) (base) ubuntu@ip-172-31-25-253:~/data_preprocessing$  
eprocessing/mecab/dept_recommendation(crawler_data)_v1.py"  
증상을 입력하세요: 무릎 관절이 아픕니다.  
추천 과: ['정형외과', '외과']  
(django) (base) ubuntu@ip-172-31-25-253:~/data_preprocessing$  
eprocessing/mecab/dept_recommendation(crawler_data)_v1.py"  
증상을 입력하세요: 설사를 하고 열이 납니다.  
추천 과: ['내과']
```

# 감사합니다

---

박장훈 송연정 양수빈 유재현 정의석