

심장병 | 고혈압 | 뇌졸중 | 당뇨

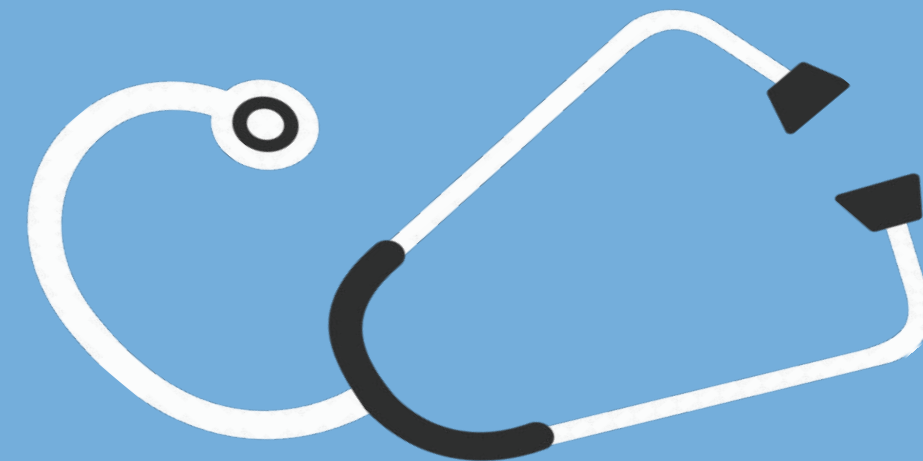


분류 분석을 통한 질병 예측

2023.05.03

김창균 | 송수린 | 송찬의 | 현정환

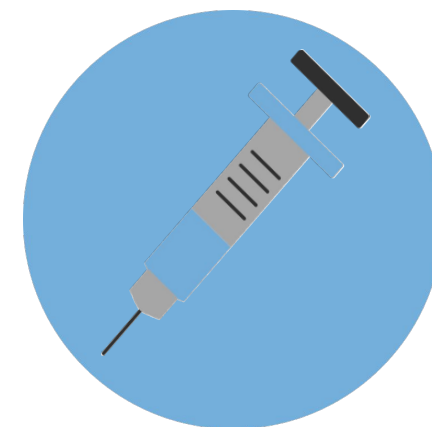
목차



1. 프로젝트 개요



2. 질병 예측 모델



3. 홈페이지 소개



4. 개선 과제 및 느낀점

프로젝트 개요



“

목적

질병을 예측하기 위한 수고로움을
덜기 위해 건강 검진 결과를 입력하여
쉽게 질병을 예측 할 수 있도록 한다.

”

프로젝트 진행 과정



질병 데이터 전처리

- 질병 유무에 따른
비대칭 데이터 처리
- 범주형 변수 처리



질병 예측 모델 구축

- 로지스틱 회귀
- 의사결정나무
- 랜덤포레스트
 - KNN
- 인공신경망
 - SVM



Flask를 이용한 홈페이지 구축

- 질병 예측 프로그램
- 건강 관련 정보
- 대학 병원 정보
- 예방 접종 및 건강 검진 정보

질병 예측 모델



심장병

변수

Age : 나이
Sex : 성별
cp : 가슴 통증 종류
trtbps : 휴식 중 혈압(mmHg)
chol : 혈중 콜레스테롤(mg/dl)
fbs : 공복 혈당
rest_ecg : 휴식 중 심전도 결과(0 정상, 1 ST-T파 이상, 2 좌심실비대)
Value: 정상/ST-T 역파 이상/가능성이 있거나 확실한 좌심실 비대
thalach : 최대 심박수
exang: 운동 유발성 협심증(0 없음, 1 있음)
oldpeak: 휴식과 관련된 운동에 의해 유발된 ST 하강
slope: 활동 ST분절 피크의 기울기(0 하강, 1 평탄, 2 상승)
ca: 형광투시법으로 확인된 주요 혈관 수
target : 심장병 여부(0 없음, 1 있음)

	언더샘플링 전	언더샘플링 후
심장병 없음	770	770
심장병 있음	784	770

모델	정확도
선형 회귀	73.6%
의사결정나무	86%
랜덤포레스트	89.3%
KNN	82%
인공신경망	89%
SVM	74.7%
DNN	79.7%

고혈압

변수

age: 나이
sex: 성별
cp: 가슴 통증 종류
trestbps: 휴식 중 혈압(mmHg)
chol: 혈중 콜레스테롤(mg/dl)
fbs: 공복 혈당
restecg: 휴식 중 심전도 결과(0 정상, 1 ST-T파 이상, 2 좌심실비대)
thalach: 최대 심박수
exang: 운동으로 인한 협심증(0 없음, 1 있음)
oldpeak: 휴식 대비 운동으로 인한 ST
slope: 활동 ST분절 피크의 기울기(0 하강, 1 평탄, 2 상승)
ca: 형광투시법으로 확인된 주요 혈관 수
thal:지중해빈혈 여부(0 결측값 1 정상, 2 고정 결함, 3 가역 결함)
target: 고혈압 진단 여부(0 없음, 1 있음)

	언더샘플링 전	언더샘플링 후
고혈압 없음	11,686	11,686
고혈압 있음	14,004	11,686

모델	정확도
선형 회귀	84.7%
의사결정나무	97.4%
랜덤포레스트	97.2%
인공신경망	99.7%
SVM	98.1%
DNN	99.6%

뇌졸중

변수

sex: 성별
age: 나이
hypertension: 고혈압 여부(0 없음, 1 있음)
heart_disease: 심장병 여부(0 없음, 1 있음)
ever_married:결혼 여부(0 미혼, 1 기혼)
wor_type: 일 경험(0 없음, 1 어릴때, 2 공무원, 3 직장인, 4 자영업)
residence_type: 사는지역(0 시골, 1 도시)
avg_glucose_level: 글리코겐 평균 수치
bmi: 비만도
smoking_status: 흡연여부(0 비흡연, 1 흡연)
stroke: 뇌졸중 진단 여부(0 없음, 1 있음)

	언더샘플링 전	언더샘플링 후
뇌졸중 없음	20,450	20,450
뇌졸중 있음	20,460	20,450

모델	정확도
선형 회귀	68.4%
의사결정나무	71.8%
랜덤포레스트	71.8%
KNN	92.4%
인공신경망	89.4%
SVM	67.8%

당뇨병

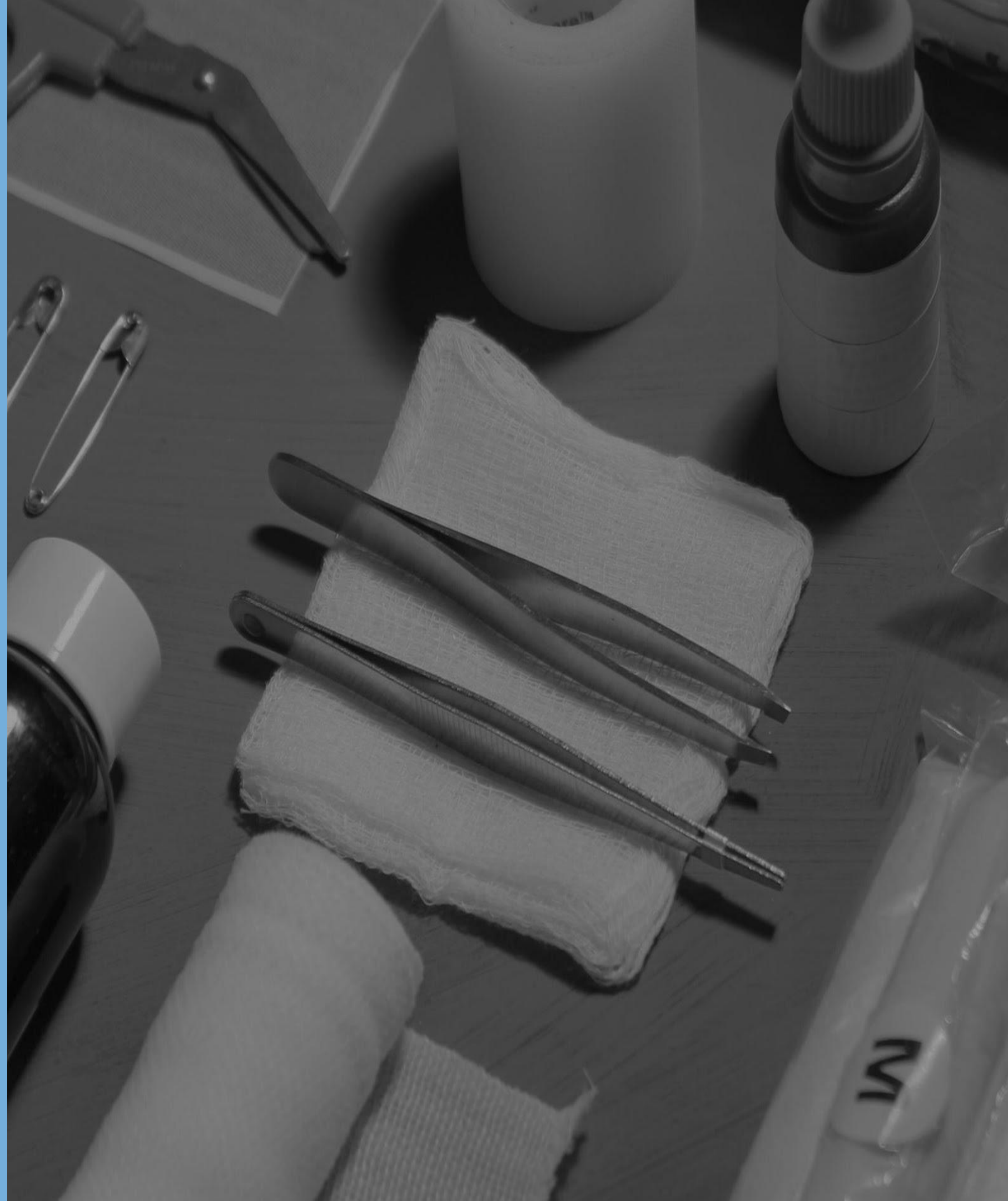
변수

Age:나이대(13분위, 1 : 18-24. 9 : 60-64, 13 : 80 이상)
Sex: 성별(1 : 남성, 0 : 여성)
Highchol: 고콜레스테롤 여부
CholCheck: 콜레스테롤 체크 여부(근 5년)
BMI: 체질량지수
Smoker: 흡연 여부(생애 100개(5갑))
HDA: 심장병/심근경색
PA: 운동 여부(근 30일)
Fruits:과일 섭취(1일 1과일) Veggies:야채 섭취(1일 1야채)
HAC: 음주 여부(남성 : 주 14회 이상 / 여성 : 주 7회 이상)
GenHlth: 전반적인 건강(5분위, 1:양호, 5:미흡)
MentHlth: 정신적 아픈 날 수/30일
PyhsHlth: 육체적 아픈 날 수/30일
DiffWalk: 걷기 어려움(계단 오르기)
Stroke: 뇌졸중 여부 High BP: 고혈압 여부 Diebetes: 당뇨 여부

	언더샘플링 전	언더샘플링 후
당뇨병 없음	213,703	35,346
당뇨병 있음	35,346	35,346

모델	정확도
선형 회귀	74.8%
의사결정나무	72.2%
랜덤포레스트	72.3%
KNN	72.7%
인공신경망	73.8%
DNN	71.6%

홈페이지 소개



홈페이지 메인

건강 예측 검사 사이트

건강 수치 기준

대학병원 정보

건강 관리 정보

질병 분석 소스

menu5



예방접종 안내



- A형 간염
- B형 간염
- 파상풍
- 인플루엔자
- 폐렴사슬알균

자세한 사항은
여기를 클릭하세요



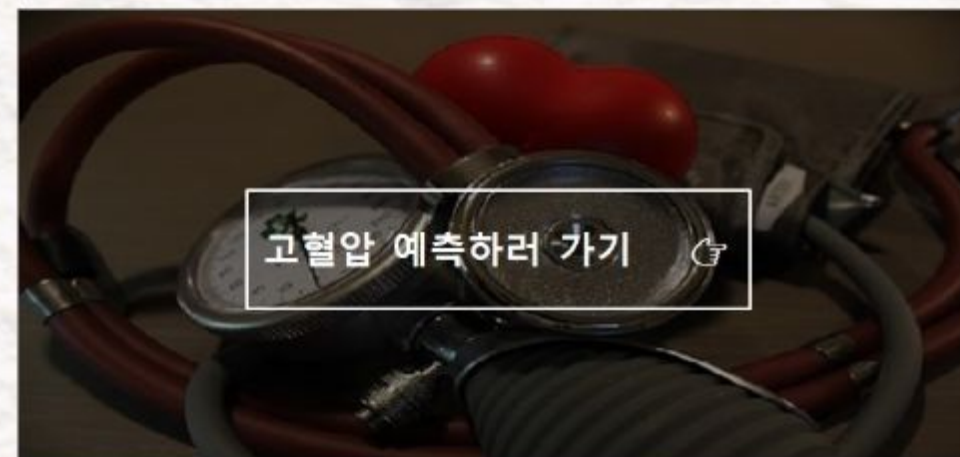
CLICK



심장병 예측하러 가기



뇌졸중 예측하러 가기



고혈압 예측하러 가기



당뇨병 예측하러 가기



건강검진 안내



2023년 일반건강검진
무료대상자

만 20세 이상
흡수년도 종생자

자세한 사항은
여기를 클릭하세요



CLICK

Tel : 02)222-2323

만든이들: 김창균, 송수린, 송찬의, 현정환

질병 예측

심장병 예측 질문지	
나이	<input type="text" value="50"/>
성별	<input type="text" value="여자"/> ▼
가슴 통증 여부	<input type="text" value="증상"/> ▼
휴식 중 혈압	<input type="text" value="80"/>
공복 혈당	<input type="text" value="100"/>
휴식 중 심전도 결과	<input type="text" value="이상"/> ▼
최대 심박수	<input type="text" value="150"/>
운동 유발성 협심증	<input type="text" value="없음"/> ▼
휴식 대비 운동으로 인한 ST 하락 폭	<input type="text" value="0.4"/>
활동 ST분절 피크의 기울기	<input type="text" value="평탄"/> ▼
형광투시법으로 확인된 주요 혈관 여부	<input type="text" value="없음"/> ▼
<div>제출</div>	

질병 예측 결과

심장마비 예측 결과표

성별: f
나이: 50.0
가슴 통증 여부: y
휴식 중 혈압: 80.0
공복 혈당: 100.0
휴식 중 심전도 결과: y
최대 심박수: 150.0
운동 유발성 협심증: n
휴식 대비 운동으로 인한 ST 하락 폭: 0.4
활동 ST분절 피크의 기울기: n
형광투시법으로 확인된 주요 혈관 여부: n

심장마비 예측 결과: 심장마비 가능성 적음(39.9%)

혈압 수치 기준

혈압 분류	수축기혈압(mmHg)		이완기혈압(mmHg)
정상혈압	<120	그리고	<80
주의혈압	120~129	그리고	<80
고혈압전단계	130~139	또는	80~89
고혈압 1기	140~159	또는	90~99
고혈압 2기	160~	또는	100~
수축기단독고혈압	140~	그리고	<90

공복 혈당 수치 기준

공복혈당분류	공복혈당수치(mg/dL)
정상수치	<100
공복혈당장애	100~125
당뇨병	126~

BMI 수치 기준

비만분류	체질량지수	
	세계기준	국내기준
저체중	18.5 미만	18.5 미만
정상	18.5 ~ 24.9	18.5 ~ 22.9
과체중	25.0 ~ 29.9	25.0 ~ 29.9
비만	30.0 이상	30.0 이상

도움이 되는 정보



건강 정보

건강 관리 정보

병명	추천 음식	주의 사항	자세히 보기
당뇨병	<ul style="list-style-type: none">• 우엉• 콩• 해조류(미역, 다시마, 해초 등)• 브로콜리	<ul style="list-style-type: none">• 규칙적인 식습관 필요• 단순당 섭취시 주의 요함• 염분 섭취 주의 요함• 동물성 기름을 식물성 기름으로 대체	당뇨병 자세히 보기
뇌졸중	<ul style="list-style-type: none">• 콩제품(두부, 비지, 유부 등)• 어패류(멸치, 새우살, 조개 등)• 해조류(미역, 다시마, 해초 등)• 속 하얀 과일(사과, 배 등)	<ul style="list-style-type: none">• 규칙적인 식습관 필요• 염분 섭취 주의 요함• 단백질 섭취 필요• 동물성 기름을 식물성 기름으로 대체	뇌졸중 자세히 보기
고혈압	<ul style="list-style-type: none">• 연어• 근대• 콩과 렌틸콩• 당근• 샐러리• 브로콜리	<ul style="list-style-type: none">• 금연 필수적으로 요함• 절주 요함• 스트레스 예방 필요	고혈압 자세히 보기
심장마비	<ul style="list-style-type: none">• 마늘• 고등어, 참치, 연어• 견과류• 유제품• 아로니아• 튀기지 않은 감자	<ul style="list-style-type: none">• 금연 요함• 절주 요함• 적절한 체중 유지 필요• 꾸준한 건강한 식사 습관 필요	심장마비 자세히 보기



대학 병원 정보

대학병원 정보

병원명	주소	전화번호	홈페이지
서울아산병원	서울특별시 송파구 올림픽로 43길 88 서울아산병원	1688-7575	서울아산병원 바로가기
삼성서울병원	서울특별시 강남구 일원로 81 삼성서울병원	1599-3114	삼성서울병원 바로가기
서울대학교병원	서울특별시 종로구 대학로 101 서울대학교병원	1588-5700	서울대학교병원 바로가기
연세세브란스병원	서울특별시 서대문구 연세로 50-1 세브란스병원	1599-1004	연세세브란스병원 바로가기
서울성모병원	서울 서초구 반포대로 222 서울성모병원	1588-1511	서울성모병원 바로가기
아주대학교병원	경기도 수원시 영동구 월드컵로 164 아주대학교병원	1688-6114	아주대학교병원 바로가기



예방접종 및 건강검진 정보

표준예방접종표

나이/조건	예방접종종류	참고사항
19~39세	A형간염	2회 접종 (0, 6 ~ 12개월)
	B형간염	항체검사결과가 음성이면 3회 접종 (0, 1, 6개월)
40~49세	파상풍-디프테리아 (Td백신)	10년마다 1회 접종, 처음 접종이면 3회 접종 (0, 1, 6 ~ 12개월)
50~64세	인플루엔자	매년 10 ~ 11월에 1회
65세 이상	폐렴사슬알균	1회
예방접종 가능한 병원찾기		

연령대별 필요한 건강검진

연령대	건강검진 항목
20대	기본적인 검사로 충분 여성의 경우에는 자궁경부암 검사 추가
30대	갑상선암 검사 필수 여성의 경우에는 유방암 검사 추가
40대	5대 암(위암, 대장암, 유방암, 자궁경부암, 간암) 검사 필수
50대	대장암, 심혈관 검사 필수 흡연자의 경우 폐암 검사 추가 남성의 경우 전립선암 검사 추가
60대 이상	각종 질환 검사 필수 노인성 질환(뇌졸중, 치매, 파킨슨병, 심혈관 질환 등) 검사 필수
자세히 알아보기	

[검진기관/병원찾기](#)



질병 분석 소스

프로젝트 정보

병명	데이터 출처	담당자	코드 정보	예측하러가기
심장병	심장병 데이터 1 심장병 데이터 2 심장병 데이터 3	송수린	<u>심장병 분석 소스</u>	<u>심장병 예측</u>
고혈압	고혈압 데이터	현정환	<u>고혈압 분석 소스</u>	<u>고혈압 예측</u>
뇌졸중	뇌졸중 데이터	송찬의	<u>뇌졸중 분석 소스1</u> <u>뇌졸중 분석 소스2</u>	<u>뇌졸중 예측</u>
당뇨병	당뇨병 데이터	김창균	<u>당뇨병 분석 소스</u>	<u>당뇨병 예측</u>





개선 과제 및 느낀점

향후 개선 과제

01

건강 검진 예약 서비스 구축

02

더욱 간단한 정보만으로
보다 정확한 질병 예측 시스템 구축

03

Django를 이용하여
완성도 높은 홈페이지 구축

느낀점

01

다중 분류로 분석하지 못한 것이
아쉬움으로 남는다.

02

다양한 모델에 대해
공부 할 수 있는 좋은 기회였다.

03

직접 프론트엔드를 구현하며
많은 것을 배울 수 있었다.

역할 분담

김창균

- 당뇨병 예측
모델 구축
- 홈페이지 구축

송수린

- 심장병 예측
모델 구축
- 홈페이지 구축

송찬의

- 뇌졸중 예측
모델 구축
- 홈페이지 구축

현정환

- 고혈압 예측
모델 구축
- 홈페이지 구축

감사합니다.