

AIB_13_Section3_Project

당뇨병 진단을 위한 머신러닝 모델기반 api 서비스
개발 및 배포

AI_13_김강호

목차

1. 서비스 소개
2. 개발 과정
3. 서비스 시연
4. 프로젝트 회고

목차

1. 서비스 소개

2. 개발 과정

3. 서비스 시연

4. 프로젝트 회고

1. 서비스 소개

- 개발 서비스 : 당뇨병 진단 목적의 머신러닝 모델 기반 Web App

- ✓ 혈압과 콜레스테롤 관련 수치, 건강상태 척도의 3가지 항목들을 이용하여 간단하게 당뇨병 여부를 진단가능

당뇨병 예측!

특성값 입력:

고혈압 여부 : 8시간 이상 공복 후 측정한 혈당이 126 mg/dL 이상인 경우는 1, 그 외는 0 선택

고 콜레스테롤 여부 : 콜레스테롤 수치가 200mg/dL 미만인 경우 0, 그 외는 1 선택

건강상태척도(주관) : 1 = 우수, 2 = 매우 양호, 3 = 양호, 4 = 보통, 5 = 나쁨

고혈압 여부(0 또는 1 입력)

고 콜레스테롤 여부(0 또는 1 입력)

건강상태척도(1, 2, 3, 4, 5 중 입력)

예측!

1. 서비스 소개

- 당뇨병 진단 Web App의 기대효과

- ✓ 의료진과 값 비싼 당뇨병 진단용 장비 없이 당뇨병 진단이 가능

- 개발도상국 등 의료여건이 부족한 사람들이 당뇨병 여부를 판단

- 당뇨병 모의 진단용으로도 사용이 가능할 것임

- 당뇨병으로 인한 사회적문제를 줄이는 데 기여

목차

1. 서비스 소개
2. 개발 과정
3. 서비스 시연
4. 프로젝트 회고

2. 개발 과정



2. 개발 과정



- CSV 데이터 수집 - 당뇨병 진단 설문조사 data

- ✓ 데이터셋 크기 : 253,680 rows X 21 columns
- ✓ 타겟 변수 : diabetes_01 (당뇨병 여부)
- ✓ 나머지 20개 column들 : 환자의 건강상태와 관련된 특성들
 - 예 : 고혈압의 여부, 고 콜레스테롤 여부, BMI, 나이

출처 : https://www.kaggle.com/datasets/alexteboul/diabetes-health-indicators-dataset?select=diabetes_012_health_indicators_BRFSS2015.csv

2. 개발 과정

1. 데이터
수집

2. 데이터
저장

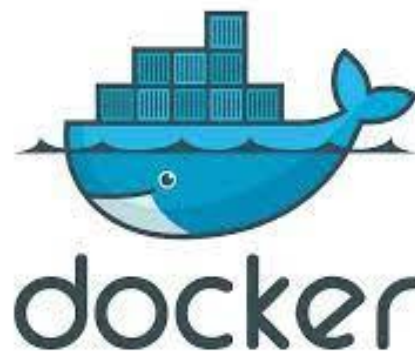
3. 데이터 분석
및 시각화

4. ML 모델링 및
web app 구축

5. web app
배포



CSV 파일을 데이터베이스[DB]에 저장



Metabase를 이용한 데이터 시각화를 위해 로컬 DB
를 Metabase 도커 컨테이너에 붙여넣기

1. 서비스 소개

2. 개발과정

3. 서비스 시연

4. 프로젝트 회고

2. 개발 과정



- Docker 컨테이너를 통해 metabase 실행
- **SQL문**으로 **DB를 가공** → 데이터 **시각화**

2. 개발 과정



- 데이터 **전처리**
- **예측모델** 생성
- 최적화



- 예측모델을 담은 **web app** 구축

2. 개발 과정



- ML 모델 기반 **web app**의 배포

목차

1. 서비스 소개
2. 개발 과정
3. 서비스 시연
4. 프로젝트 회고

3. 서비스 시연

Metabase 대시보드, 당뇨병 진단 Web App

목차

1. 서비스 소개
2. 개발 과정
3. 서비스 시연
- 4. 프로젝트 회고**

4. 프로젝트 회고

- 좋았던 점

- ✓ Section 3 내용에 대해서 더 친숙해지는 계기
- ✓ 데이터 엔지니어링에 관한 폭 넓은 지식을 공부 -> 시야 ↑
- ✓ 여러 기능적 오류들을 해결하면서 문제 해결력 향상
 - Metabase에 docker를 이용하여 데이터를 적재하는 과정
 - 머신러닝 모델을 Web app에 반영하는 과정
 - Heroku를 통해 Web app을 배포하는 과정

4. 프로젝트 회고

- 아쉬웠던 점

- ✓ 시간관리 - 에러 해결에 너무 많은 시간을 들임
- ✓ 머신러닝 모델의 성능 최적화를 많이 진행하지 못함