

AUGMENTED INTELLIGENCE LAB

2022년 1학기 업무계획

황승현

경상국립대학교 컴퓨터과학과
증강지능 연구실

01

겨울에 한 것

04

이전과 다른 점.

07

다음주 계획

02

당뇨병이란?

05

모델 설계 계획

03

개발 환경

06

개발 현황

01.

WHAT I DID
IN WINTER

01

ESD HOTDEAL

2021-07-20부터 2021-11-24까지 개발한
ESD HotDeal이 2021년 12월 29일 정식으로 개발 종료
제1회 경남소프트웨어경진대회에 출품하여
최우수상(대학총장상) 수상
제30회 소프트웨어 전시회 입상

02

DIABETES

Classification and prediction on incidence
of hypertension with blood pressure determinants in a deep learning model의 후속
연구
당뇨병 분석 모델 개발 시작

03

PID AUTO TUNING

제어계측공학과 정세교 교수님과 프로젝트
리뷰한 논문:
Neural Network-Based Self-Tuning PID Control for Underwater Vehicles

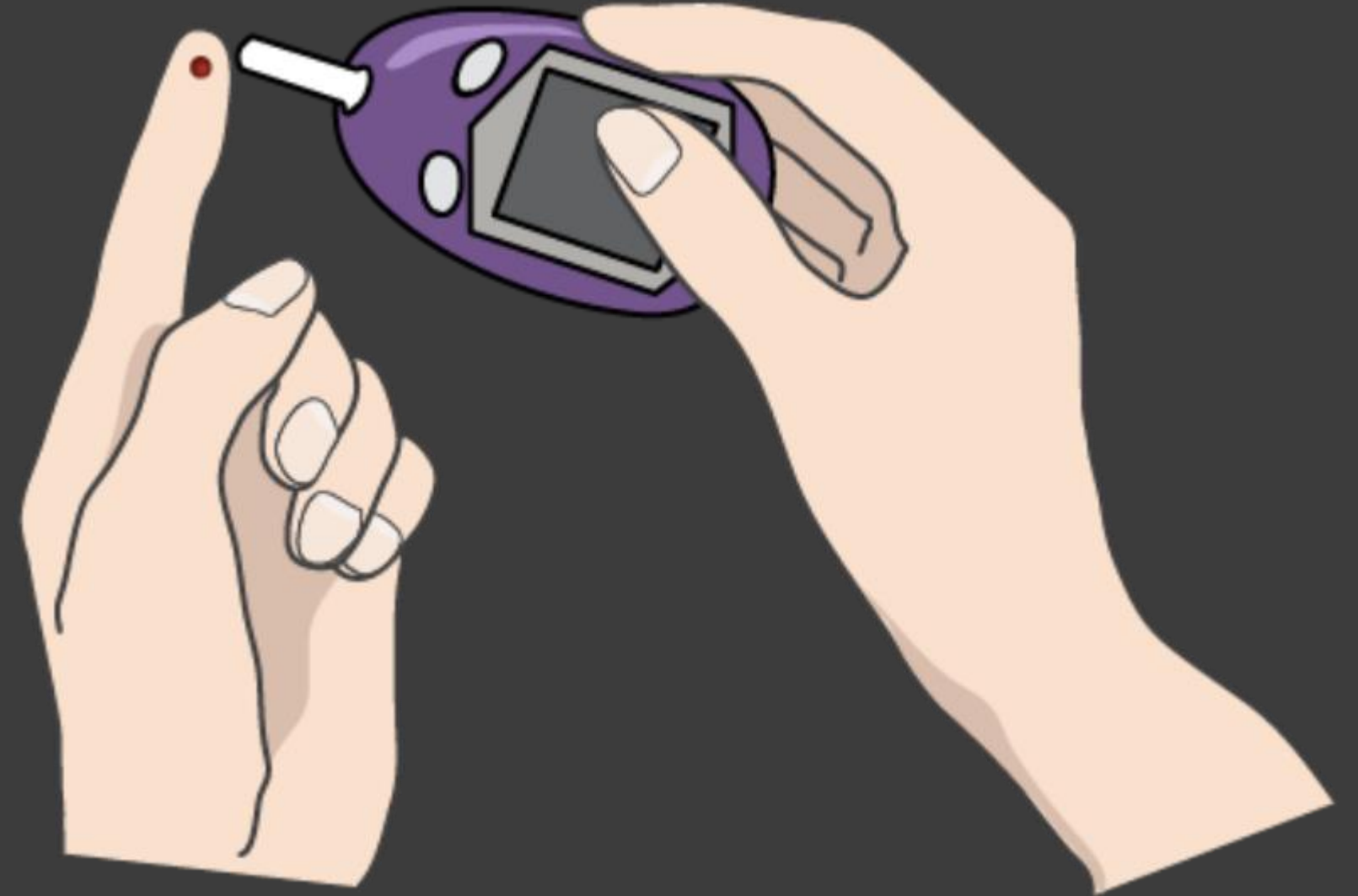
02.

WHAT IS DIABETES?

DIABETES

당뇨병이란 혈액중의 포도당(혈당)이 높아서 소변으로 포도당이 넘쳐 나오는 것. 탄수화물은 위장에서 소화효소에 의해 포도당으로 변한 다음 혈액으로 흡수됩니다. 흡수된 포도당이 우리 몸의 세포들에서 이용되기 위해서는 인슐린이라는 호르몬이 반드시 필요합니다. 인슐린은 췌장 랑게르한스섬에서 분비되어 식사 후 올라간 혈당을 낮추는 기능을 합니다. 만약 여러 가지 이유로 인하여 인슐린이 모자라거나 성능이 떨어지게 되면, 체내에 흡수된 포도당은 이용되지 못하고 혈액 속에 쌓여 소변으로 넘쳐 나오게 되며, 이런 병적인 상태를 '당뇨병' 이라고 부르고 있습니다.

출처: 대한당뇨병학회



01

DRUGINSCU

인슐린 투여 여부

1. 아니요, 2예

하나라도 해당하면 당뇨병

02

GLU0_TR

공복 혈당

126mg/dl 이상

하나라도 해당하면 당뇨병

03

GLU120_TR

당부하 120분 후 혈당

200mg/dl 이상

하나라도 해당하면 당뇨병

04

HBA1C

당화혈색소 수치

6.5% 이상

하나라도 해당하면 당뇨병

03 DEVELOPMENT _ENVIRONMENT

METHOD_1

Google Colab에서 개발
keras Sequential 모델

library	version
Python	3.7.12
tensorflow	2.7.0
numpy	1.19.5
pandas	1.19.5
matplotlib	3.2.2
google-colab	1.0.0
keras-tuner	1.1.0



METHOD_2

Google Colab에서 개발
PyTorch 모델

library	version
Python	3.7.12
tensorflow	2.7.0
numpy	1.19.5
pandas	1.19.5
matplotlib	3.2.2
google-colab	1.0.0
torch	1.10.0+cu111



04. DIFFERENCE FROM BEFORE

TUNE THE TUNER

keras tuner 파라미터 조정

Dropout rate, Learning rate 고정

unit의 개수는 이전 고혈압 분석 모델을 참고

Sequential Model의 Hidden layer의 개수도 3으로 고정.

이전 고혈압 모델에서 가장 좋은 결과.

SAW DECISION TREE

Decision Tree / Random Forest 삭제

회귀계수로 변수 중요도 판단

회귀계수란?

회귀분석에서 독립변수가 한 단위 변화함에 따라 종속변수에 미치는 영향력 크기 두 변수 사이에 상관관계가 거의 없을 때 회귀계수는 의미가 없게 된다.

당뇨병 모델에 맞게 변수의 수(column)을 줄이고, 변수의 결측값을 특성에 맞게 대체하였다. 예를 들어 AS1_HVSMAM, AS1_HVSMDU는 각각 습관적 흡연자의 하루 흡연량, 흡연 기간이다. 비흡연자는 해당 항목을 검사하지 않아서, 총 9704개 데이터에서 유효한 값(Non-Null Count)이 2370, 2390개 밖에 없었다. 이 값을 모두 0으로 대체하였다.

증강지능연구실 황승현 AILAB.GNU.AC.KR

05.

MODEL DESIGN PLAN

KERAS

keras tuner 파라미터 조정

기존 고혈압 모델에서 사용한 방법.

기존에 잘 나왔던 파라미터 그대로 이용할 것.

정확도: 88%

장점: 다루기 쉬움.

단점:

PYTORCH

PyTorch를 이용한 모델 설계

기존 모델과 다른 라이브러리 사용

PyTorch를 사용하는 이유:

현업에서 많이 사용함.

고성능, Low Level API, 속도 빠름

디버깅 용이

데이터셋이 클 때 쓰기 좋음

단점:

사용성 떨어짐(불편함)

06 . DEVELOPMENT STATUS

DEVELOPMENT STATUS

앞서 언급한 데이터 전처리, keras tuner의 파라미터 조정, Hidden Layer의 수를 모두 조정하였다. 기본적인 구조는 기존의 고혈압 모델과 동일하다.

모델을 학습했을 때, 정확도는 0.8879까지 나왔다. 고혈압 학습모델이 78정도 나온 것을 생각하면 상당히 높은 정확도이다.

trial 1

```
Trial 810 Complete [00h 00m 03s]
val_accuracy: 0.884095311164856
Best val_accuracy So Far: 0.8879587650299072
Total elapsed time: 03h 00m 10s
INFO:tensorflow:Oracle triggered exit
unit: 12 learning_rate: 0.001 dropout: 0.0.
```

trial 2

```
Trial 29 Complete [00h 00m 02s]
val_accuracy: 0.884095311164856
Best val_accuracy So Far: 0.884095311164856
Total elapsed time: 00h 01m 21s
INFO:tensorflow:Oracle triggered exit
unit: 40
```

07. NEXT WEEK'S PLAN

NEXTWEEK'S PLAN

1. 당뇨병 분석 모델 개발
 - Keras 모델 정확도 개선
 - PyTorch 모델 개발
2. 깃허브에서 PID Auto tuning 관련 소스코드 조사
 - 제어계측공학과 정세교 교수님
 - 3월 29일 세미나 예정
3. 종합설계PBL 준비
 - 주제 선정
 - 개발 범위 선정
4. 정보처리기사 필기 준비
 - 3월 5일 토요일 13시

THANK YOU

황승현

컴퓨터과학과 증강지능 연구실