



3차원 인체 스캔 데이터와 이중 에너지 엑스선 흡수기(DXA)데이터를 이용한 XGBoost기반 비만 판단 방법론 제안

Content

1. 서론

- 연구의 필요성
- related work

2. 본론

- 데이터 수집
- 연구 프레임워크
- 데이터 전처리
- 분류구성 단계, 기본특징 학습단계
- 평가지표

3. 결론

- 기대효과, 한계점

1. 서론

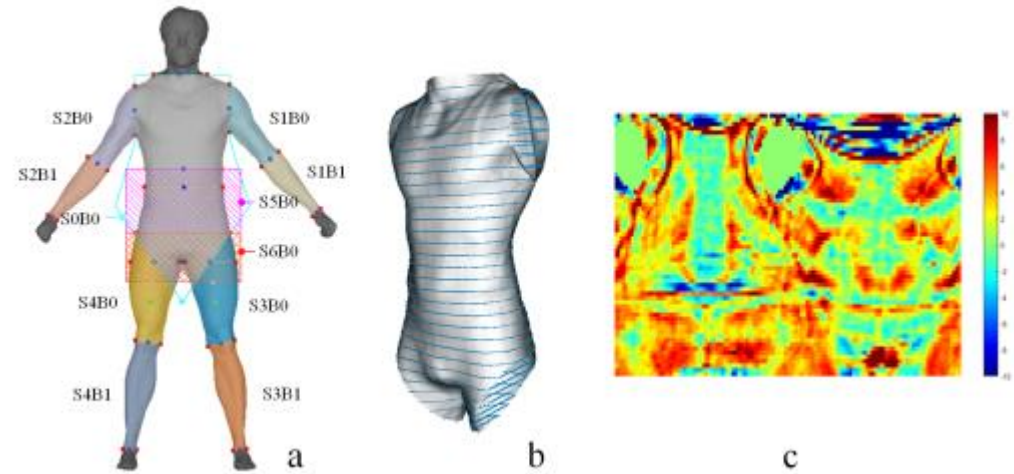
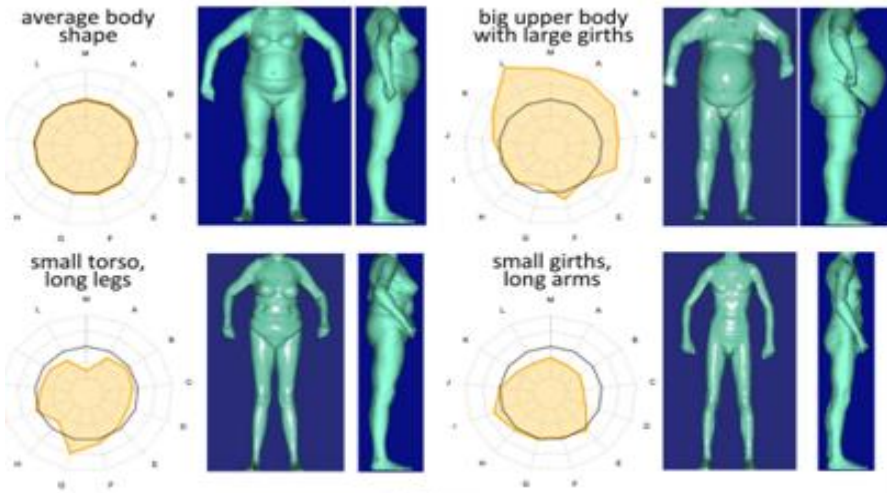
연구의 필요성

- 기존 비만 판단 지표 한계
 - 체중에 기반을 둔 BMI(Body Mass Index)는 근육과 지방을 구분하지 않음
 - 신체 부위별 정보를 제공하지 않음
- ➔ 비만이나 건강 상태를 평가하는 지수로 정확한 판단 어려움
- 기존 데이터셋의 한계
 - 기존의 3차원 인체 데이터는 각각에 대응되는 객관적인 건강 정보가 부족함
 - 체형진단과 그에 따른 헬스케어에 위한 정확한 치수와 3차원 데이터를 얻을 수 없음
 - 대부분의 데이터셋이 서양인이 피실험자임
- ➔ 한국인에 맞는 데이터셋 구축이 필요함

한국인 3차원 인체 데이터셋과 이에 대응되는 건강 정보를 수집하고, 이에 대한 연구가 필요함

1. 서론

관련 기존 연구들



3차원 스캐너로 인체 분류 혹은 비만도 판단 선행 연구

- Novel Anthropometry Based on 3DBodyscans Applied to a Large Population Based Cohort

⇒ 객관적인 의료데이터의 비교 판단 불가

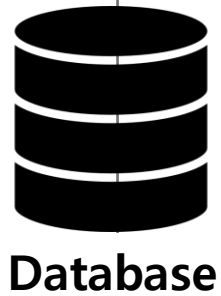
- 3D Shape-based Body Composition Prediction Model Using Machine Learning

⇒ 서양인 데이터 이고, 데이터 수가 적음, 강건성에 한계가 있음

2. 본론

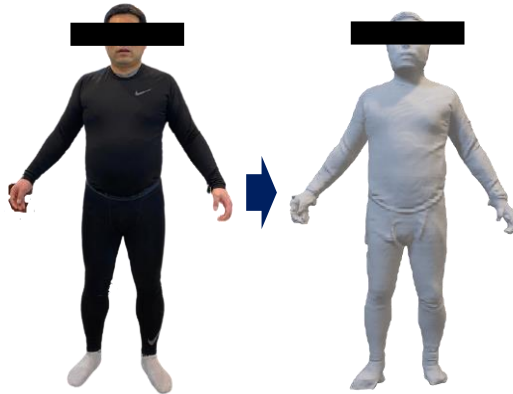
데이터 수집

- 수집 기간 : 2021-04-05 ~ 2021-04-17 (약 12일)
- 피실험자 : 62명(여성 32명, 남성 30명/20~30대)
- 수집 데이터 : 3차원 체형 스캔데이터, DXA(이중에너지엑스선흡수기) , BIA(생체 전기 임피던스 분석기)



3차원 체형 스캔데이터

- 밀착된 옷을 입고 촬영함
- 인체 길이, 둘레, 부피, 체적 등 다양한 정보 보유함



DXA

- Total Body fat%가 비만 판단의 Gold Standard로 주로 사용됨



BIA

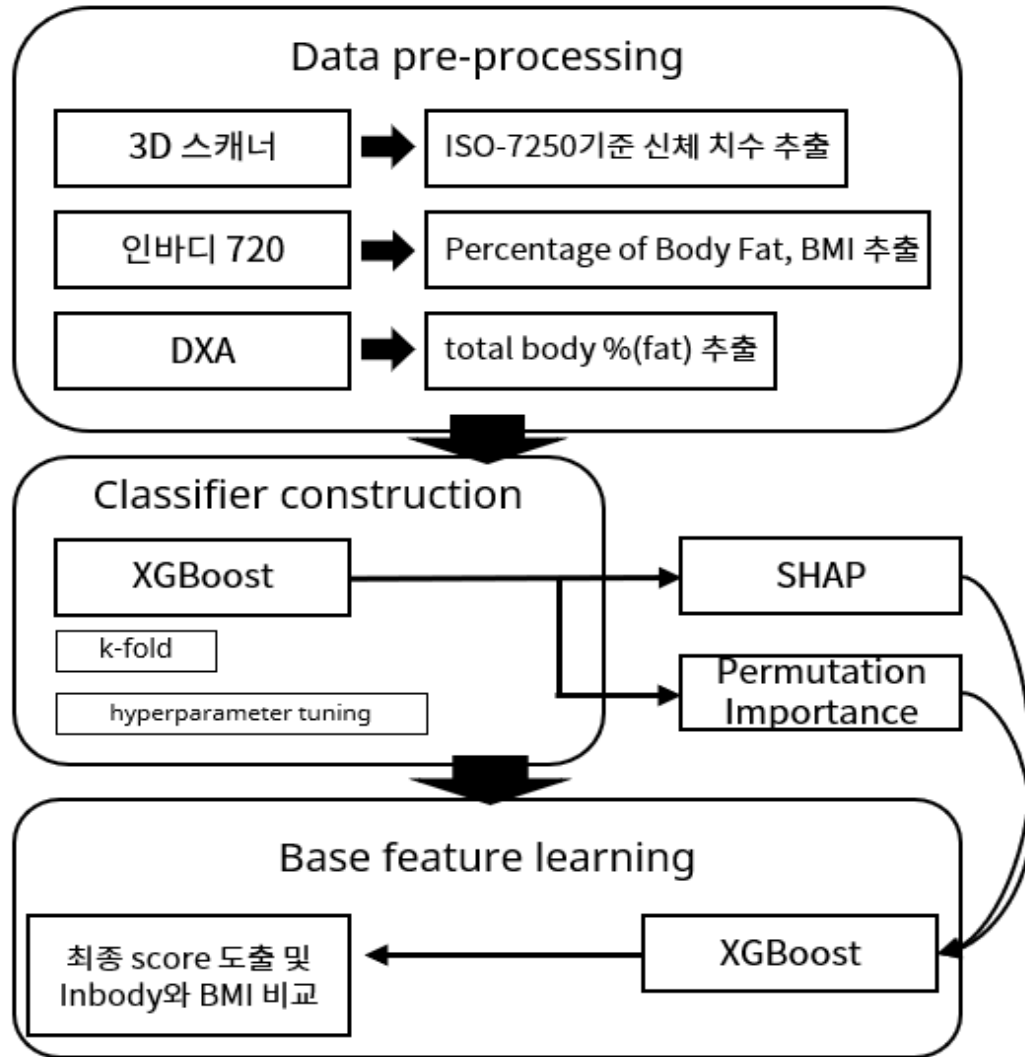
- 골격근량, 체지방량 등 체성분 평가가 가능함
- 체지방량 기준으로 비만을 판단함

체성분분석 Body Composition Analysis					
	측정치	체수분	근육량	체지방량	체중
체수분 Total Body Water	(L) 27.5 (26.3 ~ 32.1)	27.5			
단백질 Protein	(kg) 7.2 (7.0 ~ 8.6)		35.1 (33.8 ~ 41.7)	37.3 (35.8 ~ 43.7)	59.1 (43.9 ~ 59.5)
무기질 Minerals	(kg) 2.63 (2.44 ~ 2.98)				
체지방 Body Fat Mass	(kg) 21.8 (10.3 ~ 16.5)				

골격근·지방분석 Muscle-Fat Analysis					
	표준이하	표준	표준이상		
체중 Weight	55 70 85 100 115 130 145 160 175 190 205			kg	59.1
골격근량 Skeletal Muscle Mass	70 80 90 100 110 120 130 140 150 160 170			kg	19.6
체지방량 Body Fat Mass	40 60 80 100 120 140 160 180 200 220 240 260 280 300 320 340 360 380 400 420 440 460 480 500 520			kg	21.8

2. 본론

연구 프레임 워크



- 데이터 전처리

3차원 스캐너

mesh 데이터로 부터 ISO-7250 랜드마크 기준 치수 추출

BIA

인바디 720로 부터 인바디 체지방량 추출

DXA

DEXA 체지방량 추출

- 분류 구성 단계

K-fold 적용하여 XGBoost 학습

- 기본 특징 학습 단계

SHAP와 Permutation Importance를 비교하여 변수 추출 후 다시 학습

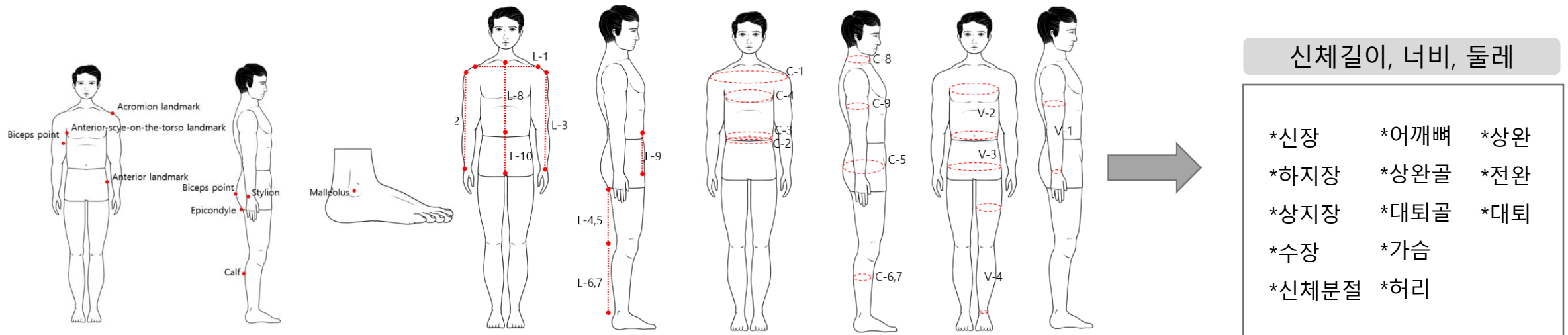
2. 본론

데이터 전처리

- 3차원 체형 스캔데이터

- ISO-7250에서 정의한 신체 랜드마크를 기준으로 60가지 항목의 신체 치수 도출

- ✓ 신체 기본정보 7개(신장, 체중, WHrR, WCR, WHR, WHtR, THR),
- ✓ 길이정보 13개(어깨, 가슴, 허리, 다리 등)
- ✓ 둘레정보 14개(배, 허리, 엉덩이 등)
- ✓ 단면적정보 13개(허리, 가슴, 엉덩이 등)
- ✓ 체적정보 8개(어깨, 상복부, 하복부 등)
- ✓ 각도정보 5개(어깨, 견갑, 허리, 엉덩이 등)



2. 본론

데이터 전처리

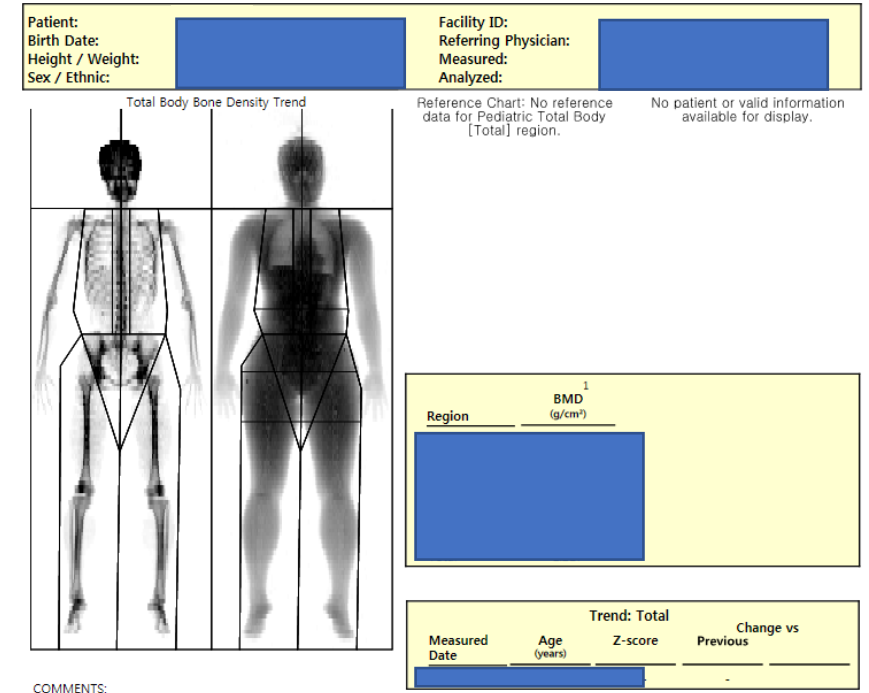
- DXA

- 방사선이 인체를 투과할 때의 투과물질의 방사선 투과율 차이로 측정함
- 체성분 분석의 Gold Standard로써 주로 사용됨
- 전신 부위별 골밀도, 면적, 지방, 제지방량 등의 데이터를 추출 가능함
- 그 중 Total body Fat(%)을 이용하여 체형을 정의하고 본 논문의 Label로 사용함

Total body Fat(%) (남성)	Total body Fat(%) (여성)	비만 판단
~ 15%	~ 18%	저체중
15% ~ 20%	18% ~ 28%	보통
20% ~ 25%	28% ~ 33%	과체중
25% ~	33% ~	비만



저체중	정상	과체중	비만
2	17	10	33



2. 본문

데이터 전처리

- BIA(생체 전기 임피던스 분석)
 - 신체의 체수분량을 측정하여 인체에 미약한 전류를 흐르게 한 뒤, 생체 전기 저항을 측정하여 체성분 분석에 이용함
 - 체수분, 단백질, 무기질, 골격근량, 체지방량 등의 신체 체성분 정보를 얻을 수 있음
 - 체지방률과 BMI를 이용한 기존 비만판단 지표와 비교

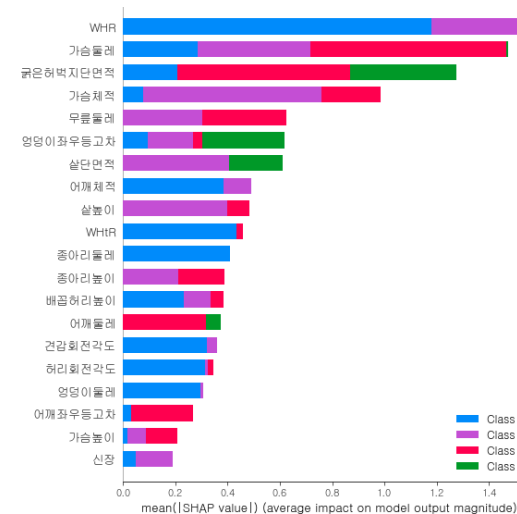
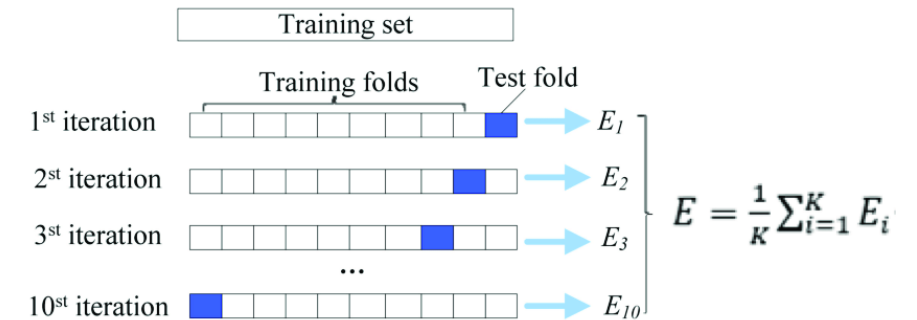
체지방율 (남성)	체지방율 (여성)	BMI	비만 판단
~ 15%	~ 18%	~ 18.5	저체중
15% ~ 20%	18% ~ 28%	18.5 ~ 23	보통
20% ~ 25%	28% ~ 33%	23 ~ 25	과체중
25% ~	33% ~	25 ~	비만



2. 본론

분류구성 단계

- 변수와 정답 값 설정
 - ISO - 7240 landmark 기준으로 추출한 60개의 치수를 변수로 설정
 - DXA 기준으로 정답 값 설정
- 클래스의 불균형 & 데이터 수 적음 => K-fold 교차검증 적용
 - 10번의 교차검증 실시 평균값을 정확도로 사용하였음
- SHAP와 Permutation importance를 통해 중요 변수 도출
 - 변수가 목적변수에 미치는 영향을 기준으로 선택
 - SHAP와 Permutation importance를 비교 분석하여 기본특징 학습 단계에서 사용할 변수 설정



Weight	Feature
0.0581 ± 0.0562	WHR
0.0484 ± 0.0204	가슴체적
0.0258 ± 0.0524	가슴둘레
0.0129 ± 0.0129	굽은허벅지단면적
0.0065 ± 0.0158	어깨체적
0.0032 ± 0.0129	종아리둘레
0 ± 0.0000	허리회전각도
0 ± 0.0000	Gender_1
0 ± 0.0000	견갑좌우등고차
0 ± 0.0000	굽은허벅지높이
0 ± 0.0000	살높이
0 ± 0.0000	신장

<Class 0, 1, 2, 3은 저체중, 정상, 과체중, 비만을 뜻함>

2. 본론

기본 특징 학습 단계

- SHAP

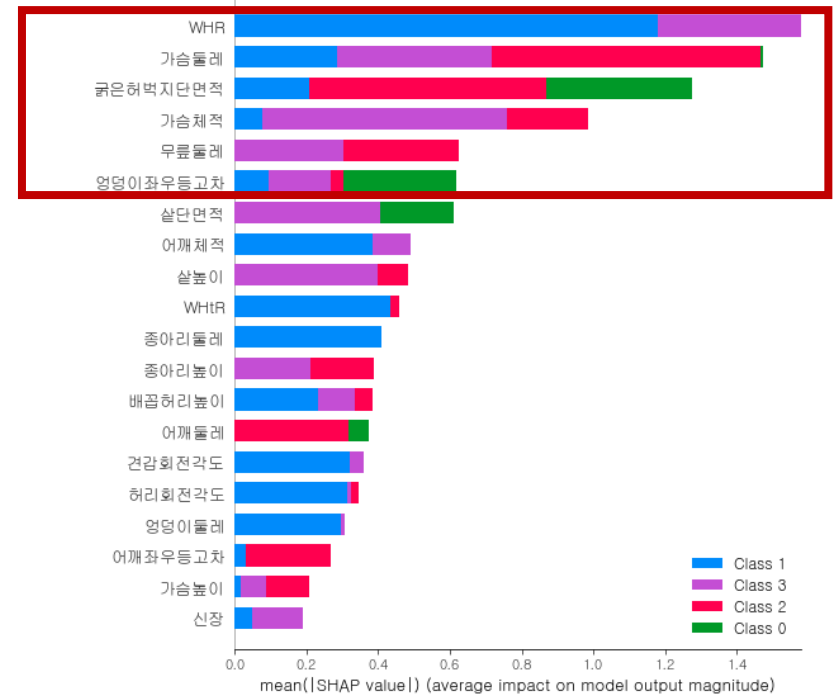
- 변수가 목적변수에 미치는 영향을 기준으로 선택

- Permutation Importance

- 특정 Feature를 쓰지 않았을 때 성능 손실에 미치는 영향 파악

→ WHR, 가슴체적, 가슴둘레, 굽은 허벅지 단면적, 어깨 체적, 종아리 둘레를
기본특징 변수로 설정

Weight	Feature
0.0581 ± 0.0562	WHR
0.0484 ± 0.0204	가슴체적
0.0258 ± 0.0524	가슴둘레
0.0129 ± 0.0129	굽은허벅지단면적
0.0065 ± 0.0158	어깨체적
0.0032 ± 0.0129	종아리둘레
0 ± 0.0000	허리회전각도
0 ± 0.0000	Gender_1
0 ± 0.0000	견갑좌우등고차
0 ± 0.0000	굽은허벅지높이
0 ± 0.0000	살높이
0 ± 0.0000	신장

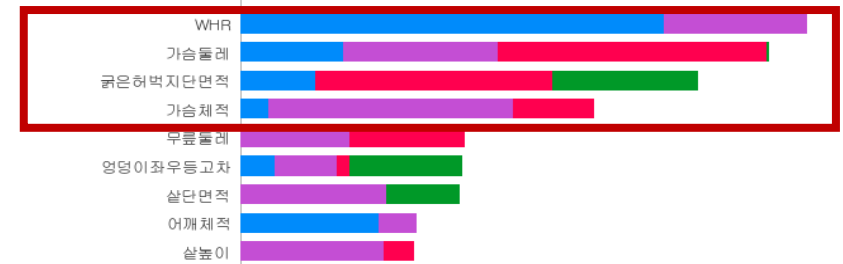


2. 본론

기본 특징 학습 단계

- 변수 추출 후 다시 학습시킨 XGBoost 모델
 - 17%포인트 증가
- 학습시킨 XGBoost 모델과 BIA, BMI 정확도 비교
 - BMI보다 19%포인트 증가
 - BIA보다 11%포인트 증가

	정확도
신체치수 XGBoost 모델	68%
기본특징 XGBoost 모델	81%
BIA (생체 전기 임피던스 분석)	70%
BMI (체질량지수)	62%



- BMI와 BIA보다 3차원 인체 데이터를 사용 한 머신러닝 방법론이 비만 판단을 내릴 때 적합 할 수 있는 가능성을 보여줌
- 비만의 정도에 따라 하나 혹은 두개의 신체 치수에 의존하는 BMI나 WHR같은 지표보다 여러 신체 치수를 반영하는 것이 비만을 판단하는 것에 유리함

3. 결론

기대 효과

- 3차원 인체 데이터와 DXA 데이터를 쌍으로 수집하여 헬스케어 데이터로서 객관적 지표가 될 수 있음
- 머신러닝 기법을 통해 데이터를 기반으로 한 새로운 비만 판단 Index 제시
- 일반적으로 수집하기 어려운 인체 데이터를 수집 및 분석함으로써 보다 실용적인 연구 방향성 제시
- 비만을 판단하는 Feature를 도출함으로써 신체의 부위별 비만 정도 및 비만의 세분화 분석 가능

한계점

- 신체를 측정하는데 필요한 3차원 스캐너와 DXA는 개인이 사용하기 어려움
- 적은 데이터로 인해 균등하게 클래스가 분포되지 않음
- DXA에서 얻을 수 있는 데이터들을 부분적으로 밖에 사용하지 않음
- Landmark를 통해 추출한 치수로 실험을 한 것이므로 3차원 데이터를 온전히 사용했다고 하기 어려움

감사합니다

Q&A