The Society of Convergence Knowledge Transactions Vol.12, No.1, pp.65-75, 2024

Research Article

pISSN : 2287-8920
https://doi.org/10.22716/sckt.2024.12.1.007

우울증 위험 척도를 예측하기 위한 우울증 선별 검사와 웨어러블 디바이스 통합 기술에 관한 연구

신탄희[1], 이대겸[2], 송상우[3], 김웅식[4]
[1,2,3]이엔유주식회사 개발부 연구원, [4]건양대학교 인공지능학과 교수

A Study on the Depression Screening and Wearable Device Integration Technology to Predict Depression Risk Scales

Tan-Hee Shin[1], Dae-Gyeom Lee[2], Sang-Woo Song[3], and Woong-sik Kim[4]
[1,2,3]Researcher, Department of Development, ENU Co.Ltd
[4]Professor, Department of Artificial Intelligence, Konyang University

[4]**Corresponding author:** wskim@konyang.ac.kr

Received February 14, 2024; **Revised** February 28, 2024; **Accepted** March 6, 2024

ABSTRACT

우울증 초기 진단 방법으로 자기보고식 우울증 선별 검사가 주로 쓰인다. 자기보고식 우울증 선별검사는 의료기관 이외에도 인터넷에서 쉽게 접할 수 있으나, 이에 비해 우울증 선별 검사 결과에 따른 적절한 사후조치가 없어 오히려 치료에 대한 동기 결여로 인해 역효과가 발생했다. 본 연구는 우울증 위험 척도 예측 결과 및 생체신호 분석 결과 시각화 제공을 통해 피검사자에게 의료기관 방문을 위한 판단 참고 정보를 제공하고자 한다. 본 연구의 우울증 선별검사와 웨어러블 디바이스 통합 기술은 데이터 취득을 위한 DAQ 시스템, 우울증 위험 척도 예측을 위한 인공지능 시스템, 인공지능모델의 입력이 될 생체신호 분석 결과와 우울증 위험 척도 예측 결과의 시각화 시스템으로 구성된다. 특히, DAQ 시스템은, 본 연구과정에서 개발한 우울증 선별 검사 통합 비침습 생체신호 취득 시나리오 기반 웨어러블 디바이스 전용 앱 포함하고 있다. 우울증 위험 척도 예측을 위한 인공지능 시스템의 경우, 직접 만 19세~50세 남,여 29명을 대상으로 우울증 선별 검사 통합 시나리오 기반 데이터를 수집을 통해 인공지능 훈련 및 평가를 수행하여 적용했다.

Self-report depression screening tests are commonly used as an initial diagnosis method for depression. However, these tests are easily accessible on the internet, but there is no appropriate follow-up measure according to the test results, which leads to a counter-effect due to the lack of motivation for treatment. This study aims to provide reference information for the decision to visit a medical institution by providing visualized results of depression risk scale prediction and biosignal analysis. The system consists of a DAQ system for data acquisition, an artificial intelligence system for depression risk scale prediction, and a visualization system for biosignal analysis results and artificial intelligence prediction results. In particular, the DAQ system includes a wearable device-specific app based on a non-invasive biosignal acquisition scenario integrated with depression screening tests developed in this study. For the artificial intelligence system, we collected data from 29 male and female subjects aged 19 to 50 based on the integrated scenario and applied the artificial intelligence model trained and evaluated through the data.

Keywords: Depression risk scale prediction, Depression screening test, Wearable devices, Biosignals, Artificial intelligence

© 2024 by The Society of Convergence Knowledge. This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. 서 론

우울증은 우울한 상태가 2주 이상 비정상적으로 길게 지속되어 일상생활에도 영향을 주는 질병으로 정신활동 능력 및 의욕 저하, 식욕 저하, 수면 장애 등이 증상으로 나타나는 질병이다. 최근 우울증 환자 수는 2017년 69만 1,164명에서 2021년 93만 3,481명으로 35.1%(연평균 7.8%) 증가했으며[1], 우울증으로 인한 사회경제적 비용은 4조 8,280억 원으로 2011년의 추정치의 3조 3,873억 원이 비해 42% 증가 한 것으로 나타났다[2].

우울증의 진단 방법으로 우울증 선별 검사, 정신건강의학과 의사와 면담, 혈액학적 검사 및 뇌 영상 검사 등이 있으며, 우울증의 초기 진단에서 일반적으로 자기보고식 우울증 선별 검사 활용을 많이 한다. 자기보고식 설문의 경우 대상자의 주관적인 경험에 따라 응답의 편향이 발생할 수 있으며[3], 일부 환자들의 경우 우울증에 대한 오해나 두려움으로 인해 우울증 선별 검사의 결과를 인정하지 않거나 치료를 기피할 가능성이 존재한다[4]. 특히 자기보고식 우울증 선별 검사는 의료기관이외에도 인터넷에서 쉽게 접할 수 있으나, 이에 비해 우울증 선별 검사 결과에 따른 적절한 사후조치가 없어 오히려 치료에 대한 동기 결여로 인해 치료에 역효과를 줄 수 있다[5].

본 연구는 Fig. 1과 같이 피검사자가 쉽게 구매할 수 있는 상용 웨어러블 디바이스(Galaxy Watch5)를 이용해 우울증 선별 검사와 비침습적 생체 신호 취득 방법을 통합하여 단시간 생체신호 데이터를 취득하고, 인공지능모델 입력이 될 생체신호 분석 결과와 우울증 위험 척도 예측 결과를 시각화하여 피검사자에게 의료기관 방문을 위한 판단 참고 정보로 제공을 통해, 의료융복합적인 서비스에 기여하고자 한다.

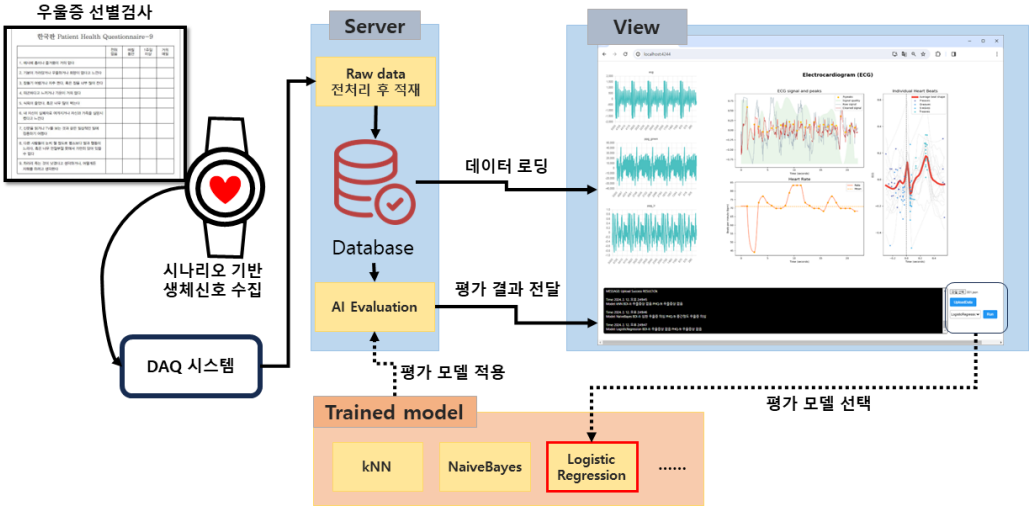


Fig. 1. The depression screening test and wearable device integrated system

2. 선행 연구

Google scholar 및 Arxiv에서 검색된 9건의 wearable device를 사용한 연구들을 조사한 결과(Table 1 참고) 참가자별 데이터 수집에 소요된 기간은 평균 9주이다. 웨어러블 기기를 착용한 상태에서 일상생활의 상호작용(통화 수, 문자 수 등), 행동 패턴(정지, 걷기, 이동 시간 등), 수면 패턴, 생체 신호(심박수, 피부 온도 등) 등을 장기간 수집하였다. 수집된 데이터를 일, 주 또는 개월 단위로 사용하여 우울증 수치를 예측하거나 우울증과 연관이 큰 바이오마커를 추출하는 것으로 나타났다. 이를 토대로 Machine Learning 모델을 훈련 및 평가를 통해 우울 증상에 대한 변화를 추정할 수 있다는 결과를 얻었다.

본 연구는 기존 연구가 1주일 이상의 장기간 일상생활로부터 데이터를 취득한 것과 달리 약 2시간 동안 우울증 선별 검사를 통합한 생체신호 취득 시나리오 기반의 ECG, PPG, SpO2, BIA 데이터를 수집하고, 인공지능 기반의 우울증 위험 척도 예측 및 인공지능 모델의 입력이 될 생체신호 분석 결과 정보 제공 시스템 피검사자에게 직접 제공한다는 차이점이 있다.

Table 1. Data collection period and model used in the preceding research

참고 문헌	수집 기간	사용 모델	설문지	사용 기기
[6]	13 주	Multiple Regression Analysis	PHQ-9	Fitbit Charge 2
[7]	1 주	Random Forest, XGBoost	DASS	Smart phone
[8]	9 주	Boosting Ensemble, Random Forest	HDRS-17	Smart phone, E4
[9]	24 주	kNN, Gradient Boosting, Random Forest, Decision Tree	PHQ-9	Smart phone, Fitbit
[10]	4 주	kNN, XGBoost, Random Forest, SVM, Logistic Regression	DASS	Smart phone, Oura
[11]	13 주	Hierarchical Logistic Regression	-	Smart phone, Fitbit
[12]	8 주	XGBoost	HAMD	Silmee W20
[13]	2 주	Gaussian Process Model	-	Shimmer 3
[14]	4~15 주	LSTM 응용 모델	PHQ-9, SGDS-K	E4
본 연구	2 시간	Logistic Regression, kNN, Naive Bayes, Random Forest, SVM 등	PHQ-9, K-BDI-2	Galaxy Watch5

3. 본 론

3.1 우울증 선별 검사 및 웨어러블 디바이스 선정

본 연구는 우울증 선별 검사에 있어, PHQ-9과 K-BDI-2를 선정하였다. 피검사자의 증상이 우울증이 아닌 스트레스로 인할 가능성을 고려하여 PSS-10(한국판 지각 스트레스 척도) 검사 또한 사용하였다. PHQ-9, K-BDI-2, PSS-10에 대한 설명은 Table 2를 참고바란다.

웨어러블 디바이스의 경우 우울증 선별검사는 의료기관 이외에도 다양한 곳에서 접할 수 있다는 점을 착안하여 상용 웨어러블 디바이스 중, 취득한 생체신호 데이터를 직접 접근 할 수 있는 Galaxy Watch5를 선정했다. Galaxy Watch5의 경우 클라우드를 경유하는 방법이 아닌 Samsung Privileged Health SDK를 통해 ECG, PPG Green, PPG IR 등의 Raw Data를 직접 취득가능하다.

Table 2. Information about selected depression screening tests

설문지	설명	점수 해석
PHQ-9[15]	PHQ-9(Patient Health Questionnaire)는 9문항 환자 건강 설문지로 일차 진료 환경에서 성인 환자들의 우울증 존재와 심각도를 판단하기 위한 진단 도구이다. 3분 이내로 작성할 수 있으며 PRIME-MD의 기존 기분 모듈을 기반으로 지난 2주간의 우울증 9가지 DSM-IV 기준으로 간단히 점수화 하였다.	0~4점: 우울 아님
		5~9점: 가벼운 우울
		10~19점: 중간 정도의 우울
		20~27점: 심한 우울
K-BDI-2[16]	K-BDI-2(Korea Beck Depression Inventory 2)는1961년 개발되어 대상자가 자신이 경험하는 우울증 증상을 평가하는 자기보고식 설문지로 지난 2주간의 수면, 식욕 저하 또는 증가를 모두 평가할 수 있도록 설계되어 있다.	0~13점: 최경도 우울
		14~19점: 경도 우울
		20~28점: 중등도 우울
		29~63점: 고도 우울

Table 2. Information about selected depression screening tests (continued)

설문지	설명	점수 해석
PSS-10[17]	PSS-10(Perceived Stress Scale)은 지난 한 달간 일상생활에서 주관적으로 느끼는 스트레스의 정도를 평가하는 척도이다.	0~12점: 정상적인 수준
		13~15점: 약간의 스트레스를 받는 상태
		16~18점: 중간 수준의 스트레스를 받는 상태
		19점 이상: 심한 스트레스를 받는 상태

3.2 우울증 선별 검사 기반 우울증 위험 척도 예측을 위한 인공지능 훈련 및 평가 데이터 수집

본 연구는 인공지능 훈련 및 평가 데이터 수집을 위해 생명윤리교육 수료 및 건양대학교 IRB(기관생명윤리위원회)로부터 승인(2023-07-016) 받은 후, Galaxy Watch5로부터 데이터 수집할 수 있는 DAQ 시스템(Data Acquisition)을 개발해 만 19세 이상 50세 미만 남녀를 모집하여 다음의 절차대로 실험자와 피실험자가 간이 칸막이로 분리된 환경에서 약 2시간 진행했다.

① Galaxy Watch 5를 착용한다. Watch의 센서 부분을 손목에 정확하게 밀착.

② Galaxy Watch 5과 DAQ System 간 통신이 연결되었는지 확인.

③ Galaxy Watch 5에 나이, 성별, 키, 몸무게를 입력하고 BIA를 측정.

④ 피검사자는 다음의 각 단계에 대해 약 30분간 각 행동을 진행. 각 단계의 시작과 끝에 커프형 혈압계로 혈압 측정.

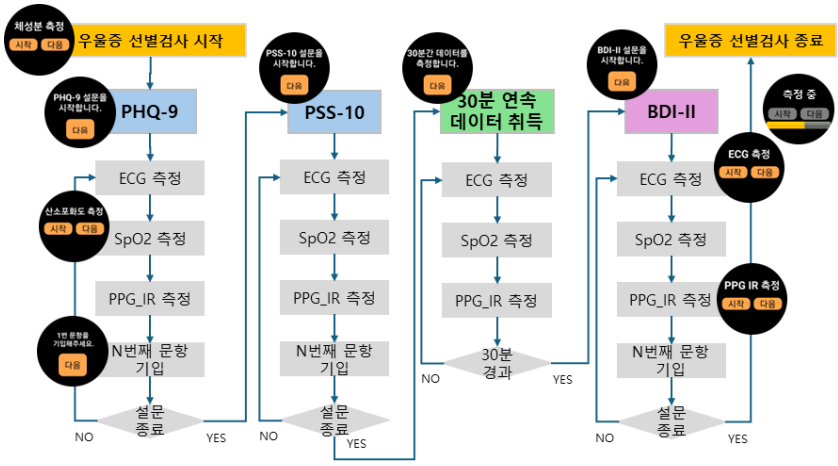


Fig. 2. A dedicated app for acquiring wearable device data based on depression screening tests

1) 우울증 선별 검사 작성(PHQ-9, PSS-10) : ECG 30초, SpO2 20초, PPG 30초를 1사이클 측정 후 한 문항 씩 기입.

2) 지속적 측정 단계: ECG 30초, SpO2 20초, PPG 30초를 지속적으로 측정.

3) 우울증 선별 검사 작성(K-BDI-2): ECG 30초, SpO2 20초, PPG 30초를 1사이클 측정 후 한 문항 씩 기입.

실험 전, 웨어러블 디바이스의 조작이 익숙하지 않은 경우를 대비해 사전 교육이 필요하며, 우울증 선별검사 시나리오 기반으로 웨어러블 디바이스 데이터 취득 전용 앱 개발(Fig. 2)하여 본 실험의 절차를 따르도록 했다. 실험의 주의사항으로 손목시계형 웨어러블 디바이스의 측성 센서와 손목 이탈이 발생할 가능성이 있다. 특히 손목이 가는 경우 이탈이 발생할 가능성이 높다. 이로 인해 모집된 검사자 29명 중, 탈락 3명을 제외하고, 총 26명의 데이터를 사용했다. 이 중, 남성 12명, 여성 14명이며, PHQ-9(>=10: 10 명, <10: 16 명), K-BDI-2(>=14: 13 명, <14: 13 명), PSS-10(>=13: 23 명, <13: 3명)으로 집계되었다.

3.3 데이터 분석

본 연구는 피검사자로부터 취득한 실시간 데이터에 대해 중점적으로 우울증 위험 척도 예측에 적용하였다. 피검사자마다 분포가 큰 차이가 없는 SpO2를 제외하고, ECG, PPG Green, PPG IR에 대해 DAQ System으로부터 추출하여, 30초, 4초 범위의 변화를 가시화하여 분석을 수행했다.

30초 범위(Fig. 3)의 ECG, PPG Green, PPG IR 신호를 분석한 결과, 측정을 시작하는 순간 발생하는 급격한 신호의 변화, 참가자의 움직임 호흡 등으로 인해 발생할 수 있는 큰 진폭과 낮은 진동수를 가지는 추세 노이즈[18] 그리고 신호 흐름에서 순간적으로 크게 벗어나는 이상치가 발견되었다. 4초간 범위(Fig. 4)의 ECG, PPG Green, PPG IR 신호를 분석한 결과, ECG 신호에서 PQRST 파형 확인이 가능하며, 심박수와 일치하는 분상 60~70회의 ECG, PPG Green, PPG IR 진동수 그리고 ECG, PPG IR에 존재하는 고주파 노이즈가 발견되었다. 특히 30초와 4초에서 공통적으로 시작지점과 끝 지점에서 큰 폭의 변화가 나타났다.

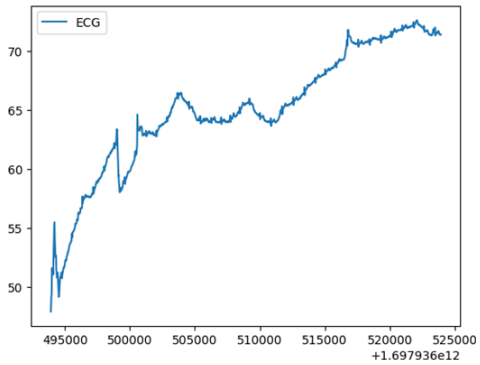


Fig. 3. Example of ECG in 30 seconds

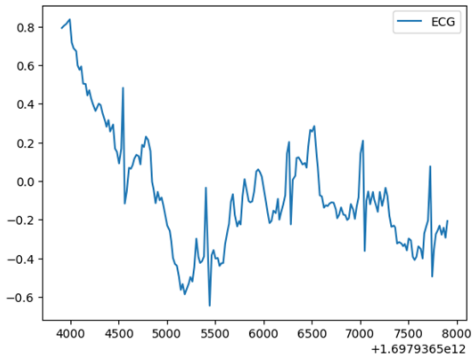


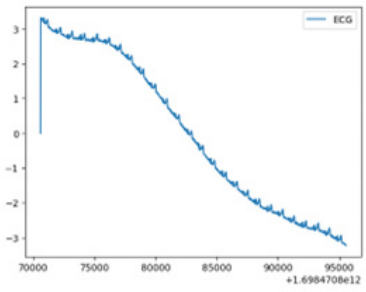
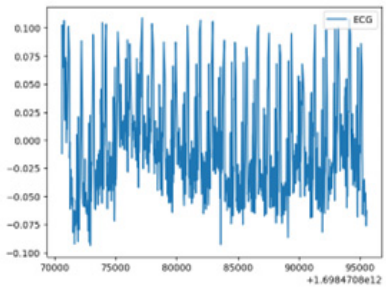
Fig. 4. Example of ECG in 4 seconds

3.4 데이터 노이즈 필터링

30초, 4초 간 ECG, PPG Green, PPG IR 데이터에서 공통적으로 발견된 시작 지점과 끝 지점의 큰 폭의 변화에 대해서는 시작 4초, 끝 1초의 데이터를 제거하여 필터링을 수행했다.

전체적으로 데이터의 기본 특성과 무관한 패턴이 표시되는 추세 노이즈에 대해서는 OLS을 사용해 추세에서 선형적인 부분을 추출해 제거하는 방법과 STL(Seasonal Trend docomposition using LOESS)을 사용해 국소적으로 가중 평균을 계산한 뒤 다항식 근사를 하는 과정을 반복하여 곡선 형태의 추세를 찾아 제거하는 방법 중, STL을 적용하여 추세 노이즈를 제거하고 데이터를 평활화 했다. Table 3을 통해 OLS가 STL에 비해 여전히 추세가 남아 있는 것을 확인 할 수 있다.

Table 3. Comparison of trend removal using OLS and STL

OLS	STL
	

이상치에 대해서는 IQR(InterQuartile Range)를 사용하여 이상치의 기준을 선정했다. IQR은 데이터를 크기 순으로 정렬하여, 상위 25% 값 Q3에서 하위 25% 값 Q1을 뺀 값으로, 이상치의 기준은 Q1-1.5*IQR 보다 작거나, Q3+1.5*IQR보다 큰 값으로 정의하였다. 이상치로 판단된 데이터에 대해서는 1차원 직선상에서 두 점의 값이 주어졌을 때, 그 사이의 값을 추정하기 위해 직선거리에 따라 선형적으로 계산(비례식)하는 선형 보간법을 적용했다.

고주파 노이즈를 포함한 각종 노이즈에 대해서는 Wavelet Transform을 적용하여 제거했다. Fourier Transform은 주어진 함수를 삼각함수의 합으로 근사해 주파수 영역으로 변환하는 반면 Wavelet Transform은 Wavelet이라하는 특정 조건을 만족시키는 함수들로 주어진 함수를 근사한다. 근사된 함수에 대해 일정 Threshold 값보다 가중치가 큰 Wavelet 함수들만 남기고, 이외에는 제거함으로써 데이터를 필터링 한다. Threshold 값을 0.01(Fig. 5), 0.02(Fig. 6), 0.05(Fig. 7), 0.1(Fig. 8)로 변환하여 원본 데이터의 훼손 효과와 노이즈 감소 효과를 경험적 관찰한 뒤, 0.05를 적용했다.

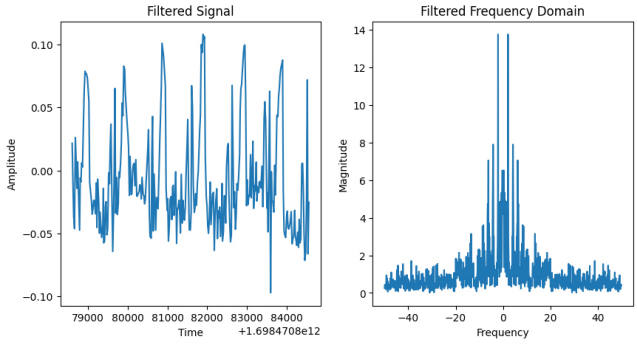


Fig. 5. Applying the wavelet transform(0.01)

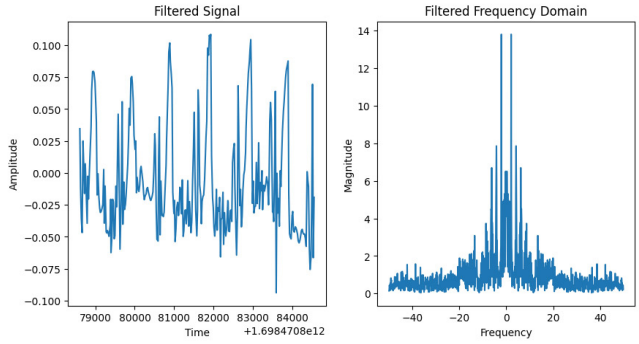


Fig. 6. Applying the wavelet transform(0.02)

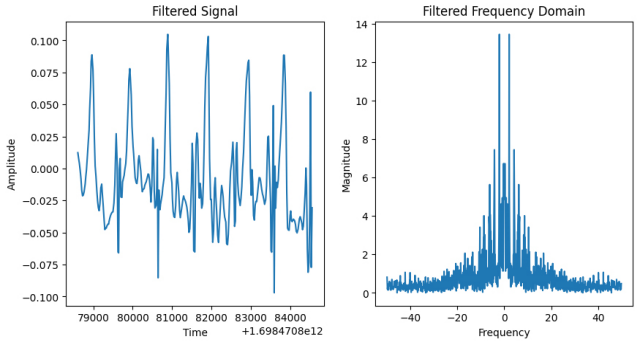


Fig. 7. Applying the wavelet transform(0.05)

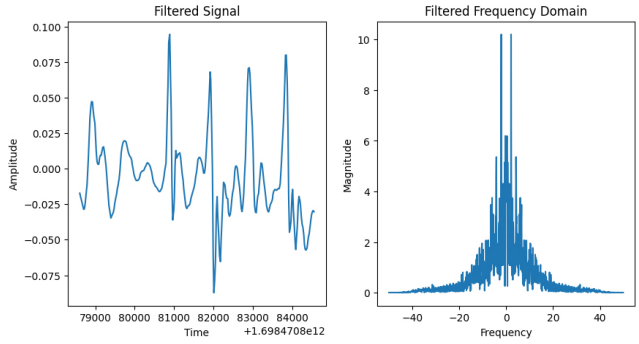


Fig. 8. Applying the wavelet transform(0.1)

3.5 데이터 Feature 추출 및 상관계수에 따른 Feature 단일화 기법 적용

피검사자 한 사람의 노이즈 필터링 된 ECG로부터 PQRST 분석을 통해 PR Interval, QT Interval, ST Segment, QRS Duration Feature를 3000개 이상의 데이터 누적 추출하였다. 이에 대해 우울증 위험 척도 예측을 위한 인공지능 훈련 및 평가에 사용하기 위한 단일 값으로 만들기 위해 평균과 표준편차를 적용했다. 노이즈 필터링 된 PPG Green, PPG IR의 경우 다른 연구[19]에서 진폭(Systolic Amplitude),너비(Pulse Width), 넓이(Pulse Area) 추출한 것을 참고하여, Fig. 9와 같이 Height, Area, Interval 25, Interval 50, Interval 75를 추출한 뒤, 평균과 표준편차를 적용했다.

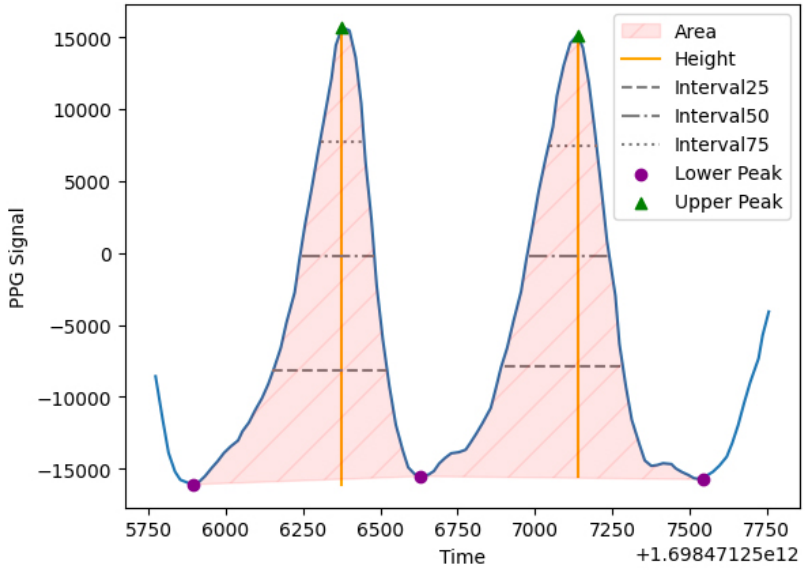


Fig. 9. PPG feature extraction

PPG Green, PPG IR로부터 추출한 Feature 간 상관계수가 0.9이상으로 높아 인공지능 모델의 성능을 낮출 수 있는 Feature에 대해 단일화를 수행했다. ECG 데이터의 경우 추출한 Feature 간 상관계수가 0.2~0.6으로 나타남에 따라 단일화를 진행하지 않았다. PPG Green과 PPG IR의 경우 이전 20개의 Feature에서 단일화 이후 8개로 줄었으며, 8개의 Feature(PPG Green Interval 50 Mean, PPG Green Interval 50 SD, PPG IR Interval 50 Mean, PPG IR Interval 50 SD, PPG Green Height Mean, PPG Green Height SD, PPG IR Height Mean, PPG IR Height SD)에 대해 Min-Max Scaling이 적용되었다.

3.6 PCA 적용을 통한 우울증 위험 척도 예측 인공지능 모델 훈련 및 평가

우울증 위험 척도 예측 인공지능 모델 훈련 및 평가를 위해, 피검사자들로부터 추출한 데이터에 대해 3 개의 Random Seed를 도입하고, 각 Random Seed에 훈련 대비 검증 비율 8:2의 k-fold 기법을 적용하여 데이터 편향의 오류 최소화를 시도하였다. 본 연구에서 시도한 모델은 Logistic Regression, kNN, Naive Bayes, Random Forest, SVM 등이 있으며, kNN의 경우 k=5로 설정하였으며, Random Forest의 경우 100개의 Tree와 Tree 최대 깊이를 2로 설정하고 Gini Impurity를 사용했다. SVM의 경우 결정 경계의 넓이에 관여하는 Gamma 값을 1/(입력된 피검사자의 사람 수 * 입력된 Feature에 대한 분산)을 사용했다.

인공지능 모델의 과적합을 방지하기 위해 고차원의 데이터를 저차원의 데이터로 사상하는 방법인 PCA를 적용하였다. PCA는 2,4,6,8 차원을 적용했다.

인공지능 평가를 위해 Accuracy, Precision, F1 Score를 사용했다. 다중 분류 문제에서는 Recall은 각 라벨마다 구한 Recall 값에 클래스 고려한 가중 평균을 적용하여 얻어 과정에서 Recall은 Accuracy와 같은 값을 가지게 된다. 따라서 본 연구에서는 Recall은 생략한다[20]. 결과 비교 기준을 최빈값으로 사용하였으며, Table 4, Table 5, Table 6에 최빈값 보다 좋은 결과가 나온 경우 파란색으로 표시되었다.

Table 4. The predicting depression risk scale PHQ-9

	Train Accuracy, 최빈값 0.45				Validation Accuracy, 최빈값 0.44			
PCA Dimension	Dim 2	Dim 4	Dim 6	Dim 8	Dim 2	Dim 4	Dim 6	Dim 8
Logistic Regression	0.49	0.63	0.70	0.77	0.40	0.3	0.25	0.26
kNN	0.52	0.62	0.57	0.55	0.32	0.48	0.48	0.49
Naive Bayes	0.51	0.69	0.74	0.83	0.22	0.39	0.23	0.26
Random Forest	0.66	0.78	0.82	0.80	0.38	0.45	0.44	0.41
SVM	0.49	0.66	0.72	0.72	0.40	0.47	0.48	0.46
	Train Precision, 최빈값 0.2				Validation Precision, 최빈값 0.22			
PCA Dimension	Dim 2	Dim 4	Dim 6	Dim 8	Dim 2	Dim 4	Dim 6	Dim 8
Logistic Regression	0.4	0.63	0.71	0.79	0.29	0.25	0.22	0.24
kNN	0.46	0.55	0.48	0.45	0.22	0.38	0.33	0.34
Naive Bayes	0.51	0.71	0.75	0.84	0.16	0.35	0.21	0.24
Random Forest	0.66	0.79	0.83	0.83	0.23	0.38	0.3	0.24
SVM	0.34	0.7	0.74	0.75	0.22	0.27	0.28	0.24
	Train F1 Score, 최빈값 0.28				Validation F1 Score, 최빈값 0.29			
PCA Dimension	Dim 2	Dim 4	Dim 6	Dim 8	Dim 2	Dim 4	Dim 6	Dim 8
Logistic Regression	0.4	0.59	0.68	0.76	0.31	0.25	0.22	0.23
kNN	0.45	0.55	0.49	0.46	0.25	0.39	0.38	0.38
Naive Bayes	0.48	0.69	0.73	0.83	0.17	0.34	0.2	0.22
Random Forest	0.59	0.73	0.79	0.79	0.25	0.37	0.34	0.29
SVM	0.35	0.6	0.67	0.67	0.28	0.33	0.34	0.31

Table 5. The predicting depression risk scale K-BDI-2

	Train Accuracy, 최빈값 0.52				Validation Accuracy, 최빈값 0.51			
PCA Dimension	Dim 2	Dim 4	Dim 6	Dim 8	Dim 2	Dim 4	Dim 6	Dim 8
Logistic Regression	0.57	0.75	0.79	0.82	0.41	0.5	0.46	0.46
kNN	0.56	0.59	0.68	0.64	0.36	0.49	0.63	0.53
Naive Bayes	0.59	0.77	0.82	0.86	0.39	0.52	0.41	0.42
Random Forest	0.70	0.80	0.84	0.84	0.40	0.56	0.58	0.49
SVM	0.56	0.77	0.82	0.82	0.42	0.54	0.56	0.53
	Train Precision, 최빈값 0.27				Validation Precision, 최빈값 0.28			
PCA Dimension	Dim 2	Dim 4	Dim 6	Dim 8	Dim 2	Dim 4	Dim 6	Dim 8
Logistic Regression	0.47	0.76	0.78	0.84	0.27	0.51	0.53	0.54
kNN	0.51	0.64	0.55	0.53	0.29	0.43	0.54	0.45
Naive Bayes	0.6	0.79	0.84	0.88	0.38	0.47	0.39	0.36
Random Forest	0.7	0.8	0.83	0.86	0.29	0.55	0.49	0.42
SVM	0.43	0.83	0.85	0.86	0.27	0.4	0.38	0.34
	Train F1 Score, 최빈값 0.28				Validation F1 Score, 최빈값 0.29			
PCA Dimension	Dim 2	Dim 4	Dim 6	Dim 8	Dim 2	Dim 4	Dim 6	Dim 8
Logistic Regression	0.47	0.74	0.78	0.82	0.31	0.47	0.45	0.43
kNN	0.51	0.64	0.6	0.57	0.29	0.42	0.55	0.46
Naive Bayes	0.55	0.76	0.81	0.86	0.35	0.46	0.38	0.37
Random Forest	0.66	0.77	0.82	0.84	0.32	0.5	0.47	0.41
SVM	0.44	0.74	0.8	0.8	0.31	0.44	0.44	0.41

Table 6. The predicting stress risk scale PSS-10

	Train Accuracy, 최빈값 0.54				Validation Accuracy, 최빈값 0.36			
PCA Dimension	Dim 2	Dim 4	Dim 6	Dim 8	Dim 2	Dim 4	Dim 6	Dim 8
Logistic Regression	0.57	0.72	0.77	0.79	0.25	0.39	0.39	0.36
kNN	0.61	0.67	0.67	0.64	0.25	0.34	0.39	0.33
Naive Bayes	0.60	0.73	0.75	0.80	0.26	0.33	0.34	0.33
Random Forest	0.85	0.90	0.95	0.97	0.16	0.28	0.27	0.21
SVM	0.63	0.81	0.84	0.86	0.19	0.27	0.24	0.25
	Train Precision, 최빈값 0.29				Validation Precision, 최빈값 0.15			
PCA Dimension	Dim 2	Dim 4	Dim 6	Dim 8	Dim 2	Dim 4	Dim 6	Dim 8
Logistic Regression	0.57	0.72	0.78	0.79	0.16	0.42	0.43	0.4
kNN	0.59	0.67	0.67	0.64	0.19	0.4	0.51	0.4
Naive Bayes	0.6	0.73	0.76	0.8	0.28	0.4	0.44	0.43
Random Forest	0.88	0.92	0.95	0.97	0.18	0.22	0.27	0.2
SVM	0.66	0.81	0.84	0.89	0.1	0.29	0.25	0.27
	Train F1 Score, 최빈값 0.35				Validation F1 Score, 최빈값 0.36			
PCA Dimension	Dim 2	Dim 4	Dim 6	Dim 8	Dim 2	Dim 4	Dim 6	Dim 8
Logistic Regression	0.53	0.72	0.77	0.78	0.18	0.37	0.38	0.36
kNN	0.58	0.65	0.66	0.63	0.21	0.32	0.38	0.31
Naive Bayes	0.9	0.72	0.74	0.79	0.23	0.32	0.35	0.34
Random Forest	0.86	0.9	0.96	0.96	0.14	0.24	0.23	0.18
SVM	0.57	0.8	0.83	0.85	0.13	0.25	0.21	0.22

3.7 생체신호 분석 결과 및 우울증 위험 척도 예측 결과 시각화 프로그램 개발

본 연구는 피검사자에게 의료기관 방문을 위한 판단 참고 정보를 제공하기 위해 인공지능 모델 입력이 될 생체신호 분석 결과와 우울증 위험 척도 예측 결과를 시각화하는 프로그램을 개발했다. Fig. 10의 ①, ②, ③은 노이즈 필터링된 ECG, PPG Green,

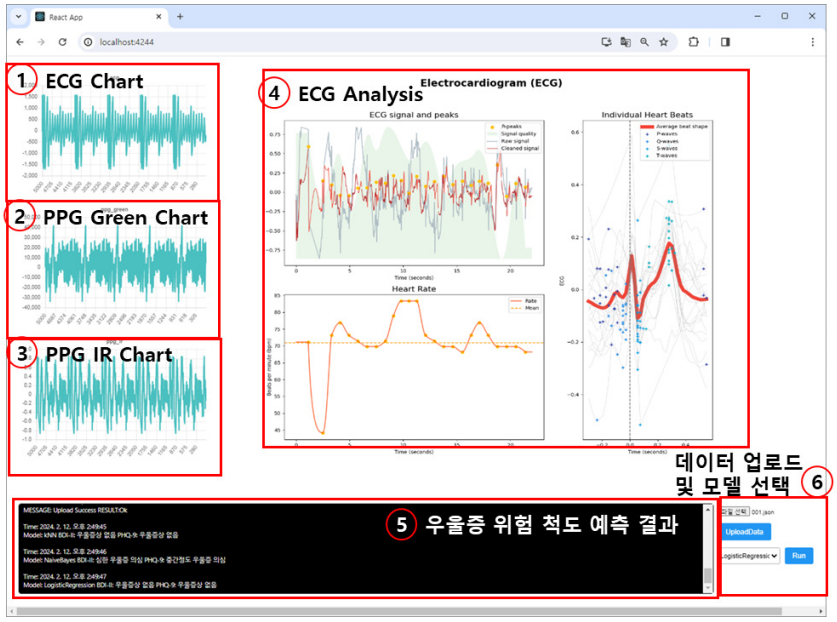


Fig. 10. Depression risk scale prediction and biometric signal analysis visualization program

PPG IR에 대한 Chart를 보여주며, 피검사자로부터 추출된 생체신호 가시화에 사용된다. ④은 ECG 데이터의 PQRST 분석에 따른 자세한 결과를 Chart로 표현된다. 이 결과가 우울증 위험 척도 예측을 위한 인공지능 시스템의 직접적인 입력 값이 된다. ⑥의 경우 사용자가 ⑤에서 선택한 인공지능 모델이 예측한 결과를 보여준다.

본 프로그램 Node.js와 React.js, SQLite, Python을 이용해 개발되었으며, 향후 피검사자에게 우울증 위험 척도 예측 과정을 설명하고, 결과를 해석할 수 있는 자료를 추가하여 결과에 대한 신뢰를 높이도록 연구 개발이 필요하다.

4. 결 론

본 연구는 우울증 위험 척도 예측을 위해 우울증 선별 검사와 통합된 비침습 생체신호 취득을 위한 시나리오 기반 웨어러블 디바이스 전용앱이 포함된 데이터 취득 시스템, 인공지능 모델의 입력이 될 생체신호 분석 결과와 우울증 위험 척도 예측 결과 시각화 프로그램을 개발했다. 그리고 우울증 위험 척도 예측을 위한 인공지능 훈련 및 평가를 위해 만 19세~50세 남, 여 29명을 모집하여, 중 탈락 3명을 제외한, 총 26명으로부터 생체신호(ECG, PPG Green, PPG IR 등)과 우울증 선별 검사(PHQ-9, K-BDI-2) 결과를 수집했다. 수집된 생체 신호 경우 다양한 노이즈(추세, 이상치, 고주파 노이즈 등)가 포함되어 있어, STL, IQR 그리고 Wavelet Transform 노이즈 필터링을 적용했다. 이후, ECG와 PPG Green, PPG IR로부터 Feature를 추출 및 상관관계 0.9이상인 Feature에 대해 단일화 적용했다. 인공지능 모델의 과적합 방지를 위해 PCA 기법을 도입하여 훈련 및 평가를 실시했다. 실시한 결과 PCA를 통해 6차원으로 변한 공간에서 훈련된 kNN, SVM 모델이 Accuracy, Precision, F1 Score 지표에서 최빈값을 상화하는 것을 확인했다. 특히 kNN 모델은 K-BDI-2 검사 지표에서 모두 0.5 이상의 결과를 기록했다. 이에 따라, 단시간 측정한 생체신호가 우울증 선별검사 결과 기반의 우울증 위험 척도 예측에 낮은 수준의 유의미한 영향을 준다는 사실을 확인했다. 본 연구가 평균 9주를 동안 피검사자의 생활 패턴과 수면패턴 등을 다각도로 데이터를 수집한 선행연구와 다르게 , 실험실 환경에서 약 2시간 동안 웨어러블 디바이스 전용 앱을 활용해 생체 신호와 우울증 선별 검사를 데이터를 수집한 상황에 대한 고려가 필요하며, 향후 선행 연구의 방법론을 도입한 확대 연구가 필요하다.

Acknowledgement

※본 연구는 중소벤처기업진흥공단에서 주관하는 2023년도 중소기업 계약학과 산학협력 프로젝트를 통해 수행된 연구임.

References

1. 건강보험심사평가원, “최근 5년(2017~2021년) 우울증 과 불안장애 진료현황 분석”, 2022.
2. 이진규, 선우성문, 최문정, 김동하, 박규민, 고준석, 장성만, “2019년 국내 우울증의 사회경제적 비용 산출”, 생물치료정신의학, 제27권 제3호, pp. 237-244, 2021.
3. 조민규, 류성진, “자기보고식 설문조사에서 응답자의 지각 및 응답편향에 관한 연구: 인터넷게임이용시간을 중심으로”, 언론과학연구, 제16권 제4호, pp. 335-373, 2016.
4. 박종익, 전미나, “정신질환에 대한 사회적 편견”, 신경정신의학, 제55권 제4호, pp. 299-309, 2016.
5. 정현숙, 임현우, 조선진, 김철민, 정소나, 송후림, 정영은, 고기동, “일차의료에서 우울증의 선별과 사후관리”, 가정의학회지, 제2권 1호, pp. 15-23, 2012.
6. Y. Rykov, T. Thach, I. Bojic, G. Christopoulos, and J. Car, “Digital Biomarkers for Depression Screening With Wearable Devices: Cross-sectional Study With Machine Learning Modeling”, JMIR mHealth and uHealth, Vol. 9, No. 10, e24872, 2021.

7. N.C. Jacobson, and Y.J. Chung, “Passive Sensing of Prediction of Moment-To-Moment Depressed Mood among Undergraduates with Clinical Levels of Depression Sample Using Smartphones”, Sensors (Basel), Vol. 20, No. 12, 3572, 2020.
8. P. Pedrelli, et al., “Monitoring Changes in Depression Severity Using Wearable and Mobile Sensors”, Frontiers in Psychiatry, Vol. 11, 584711, 2020.
9. T. Mullick, A. Radovic, S. Shaaban, and A. Doryab, “Predicting Depression in Adolescents Using Mobile and Wearable Sensors: Multimodal Machine Learning-Based Exploratory Study”, JMIR Formative Research, Vol. 6, No. 6, e35807, 2022.
10. K.O. Asare, I. Moshe, et al., “Mood ratings and digital biomarkers from smartphone and wearable data differentiates and predicts depression status: A longitudinal data analysis”, Pervasive and Mobile Computing, Vol. 83, 101621, 2022.
11. A. Horwitz, E. Czyz, N. Al-Dajani, et al., “Utilizing daily mood diaries and wearable sensor data to predict depression and suicidal ideation among medical interns”, Journal of Affective Disorders Vol. 313, pp. 1-7, 2022.
12. Y. Tazawa, K.C. Liang, M. Yoshimura, et al., “Evaluating depression with multimodal wristband-type wearable device: screening and assessing patient severity utilizing machine-learning”, Heliyon, Vol. 6, No. 2, e03274, 2020.
13. I. Peis, J.D. López-Moríñigo, M.M. Pérez-Rodríguez, et al., “Actigraphic recording of motor activity in depressed inpatients: a novel computational approach to prediction of clinical course and hospital discharge”, Scientific Reports, Vol.10, 17286, 2020.
14. J. Choi, S. Lee, S. Kim, D. Kim, and H. Kim, “Depressed Mood Prediction of Elderly People with a Wearable Band”, Sensors (Basel), Vol. 22, No. 11, 4174, 2022.
15. 박승진, 최혜라, 최지혜, 김건우, 홍진표, “한글판 우울증 선별도구(Patient Health Questionnaire-9, PHQ-9)의 신뢰도와 타당도”, 대한불안의학회지, 제6권 제2호, pp. 119-124, 2010.
16. 김지혜, 황순택, 홍상황, 이은호, “Beck Depression Inventory-II”, 한국심리주식회사, 2014.
17. 박준호, 서영석, “대학생을 대상으로 한 한국판 지각된 스트레스 척도 타당화 연구”, 한국심리학회지, 제29권 제3호 pp. 611-629, 2010.
18. S. Chatterjee, R.S. Thakur, R.N. Yadav, L. Gupta, and D.K. Raghuvanshi, “Review of noise removal techniques in ECG signals”, IET Signal Process, Vol. 14, No. 9, pp. 569-590, 2020.
19. J. Park, H.S. Seok, S.S. Kim, and H. Shin, “Photoplethysmogram Analysis and Applications: An Integrative Review”, Frontiers in Physiology, Vol. 12, 808451, 2022.
20. https://www.evidentlyai.com/classification-metrics/multi-class-metrics