

원본 논문

우울증 진단 및 심각도 평가를 위한 설명 가능한 AI 활동 데이터에 기반한 분류: 개발 및 해석 가능한 프레임워크에 대한 평가 연구

이프티카르 아흐메드¹, 박사; Anushree Brahmacharimayum¹, MA; 라자 하심 알리¹
무함마드 오와이스 아흐마드², 박사

탈하 알리 칸 박사¹ 박사;

¹독일 포츠담 유럽응용과학대학교 소프트웨어공학과
²칼스타드 대학교 수학 및 컴퓨터 과학과, 칼스타드, 스웨덴

교신 저자:

Muhammad Ovais Ahmad, PhD 수학과 컴퓨

터 과학 Karlstad University Universitetsgatan 2 Karlstad 65188 스웨덴

전화: 46 76-113 22 49 이메일

일: ovais.ahmad@kau.se

추상적인

배경: 우울증은 전 세계적으로 가장 흔한 정신 질환 중 하나로, 약 2억 8천만 명에게 영향을 미치며 진단되지 않거나 오진되는 경우가 많습니다. 웨어러블 기기의 보급이 확대됨에 따라 활동 수준을 지속적으로 모니터링할 수 있게 되었고, 이는 데이터 기반의 우울증 조기 발견 및 심각도 평가를 위한 새로운 가능성을 제시합니다.

하지만 기존의 기계 학습 모델은 우울증의 중복되는 하위 유형을 구분하는 데 있어 성능이 떨어지는 경우가 많고, 임상적 수용에 필수적인 요소인 설명력이 부족한 경우가 빈번합니다.

목표: 본 연구는 웨어러블 활동량 측정 데이터를 활용하여 우울증을 감지하고 그 심각도를 분류하는 해석 가능한 기계 학습 프레임워크를 개발 및 평가하는 것을 목표로 하며, 동시에 불균형 데이터셋 및 제한된 모델 투명성과 같은 일반적인 문제점을 해결하고자 한다.

방법: 본 연구에서는 Depresjon 데이터셋을 사용하고, 클래스 불균형을 완화하기 위해 ADASYN(Adaptive Synthetic Sampling)을 적용했습니다. 원시 활동 데이터에서 여러 통계적 특징(예: 전력 스펙트럼 밀도 평균 및 자기상관)과 인구통계학적 속성(예: 연령)을 추출했습니다. 로지스틱 회귀, 서포트 벡터 머신, 랜덤 포레스트, XGBoost 및 신경망의 다섯 가지 머신러닝 알고리즘을 정확도, 정밀도, 재현율, F1 점수, 특이도 및 매튜 상관계수를 통해 평가했습니다. 또한 예측 요인을 규명하기 위해 Shapley Additive Explanations(SHAP)과 Local Interpretable Model-agnostic Explanations(LIME)을 사용했습니다.

결과: XGBoost는 이진 분류에서 84.94%, 다중 클래스 심각도 분류에서 85.91%의 가장 높은 전체 정확도를 달성했습니다. SHAP과 LIME은 전력 스펙트럼 밀도 평균, 연령 및 자기상관을 주요 예측 변수로 밝혀냈으며, 이는 일주기 리듬 장애가 우울증에 미치는 영향을 강조합니다.

결론: 본 연구에서 제시하는 해석 가능한 프레임워크는 우울증 환자와 비우울증 환자를 안정적으로 구분하고, 경증 우울증과 중등도 우울증을 구별합니다. SHAP 및 LIME의 포함은 투명하고 임상적으로 의미 있는 통찰력을 제공하며, 설명 가능한 인공지능이 정신 건강 관리에서 조기 발견 및 개입 전략을 강화할 수 있는 잠재력을 강조합니다.

JMIR 정신 건강 2025;12:e72038; doi: [10.2196/72038](https://doi.org/10.2196/72038)

핵심어: 인공지능; 설명 가능한 인공지능; 우울증; 정신 건강; 머신러닝; 활동 데이터

소개

정신 질환은 전 세계적으로 10억 명이 넘는 사람들이 다양한 유형의 정신 질환으로 고통받는 세계적인 현상입니다.

정신 질환은 일반적으로 개인의 사고, 대인 관계 및 감정에 영향을 미칩니다. 적절한 시기에 진단 및 치료를 받지 못하면 다른 질병의 위험 증가 및 자살과 같은 심각한 결과를 초래할 수 있습니다. 우울증은 개인의 일상생활과 행복에 부정적인 영향을 미치는 가장 흔한 정신 질환 중 하나입니다. 연구에 따르면 코로나19 팬데믹 기간 및 이후에 우울증과 불안 장애 발생률이 증가했으며, 이는 남녀 모두에게 영향을 미치는 것으로 나타났습니다 [1,2]. 젊은 층이 노년층보다 정신 질환으로 더 많이 고통받는 것으로 밝혀졌습니다 [3]. 전 세계적으로 약 2억 8천만 명이 우울증을 앓고 있는 것으로 추산됩니다 [4].

우울증은 성격 변화나 어떤 상황이나 환경에 대해서는 갑자기 히스테리적인 태도를 보이는 것이 특징입니다.

우울증은 두 번째로 흔한 사망 원인으로 여겨지며, 80만 명이 우울증으로 사망하는 것으로 추산됩니다 [5]. 우울증은 삶에 영향을 미치는 절망감, 비관주의, 무쾌감증을 유발하는 정서 장애입니다. 이는 일상생활의 모든 측면에서 활동을 방해하고 주변 사람들에게도 영향을 미칩니다.

이러한 유형의 정신 질환의 원인은 환경적 요인과 유전적 요인에서 비롯됩니다 [6]. 우울증은 심각도에 따라 경증, 중등증, 중증으로 분류될 수 있습니다 [7]. 의사들이 초기 단계에서 이를 진단하는 것은 어렵고, 비타민 결핍증과 같은 질환으로 오진되는 경우가 많습니다 [8]. 또한, 우울증은 암 [9] 및 심혈관 질환 [10]의 위험 증가와도 연관되어 있으며, 이는 개인의 수명을 단축시킬 가능성이 있습니다.

높은 유병률을 보임에도 불구하고 많은 경우 치료되지 않은 채로 방치되는 경우가 많습니다 [11].

우울증 평가 척도는 우울증의 심각도 단계를 감지하는 데 사용되는 설문지를 포함하는 선별 도구입니다 [12]. Montgomery-Asberg 우울증 평가 척도(MADRS)는 Marie Asberg가 설계했습니다 [13].

이 척도는 명백한 슬픔, 보고된 슬픔, 내적 긴장, 수면 감소, 식욕 감소, 집중력 저하, 무기력, 무감각, 비관적 사고, 자살 충동을 포함한 우울증의 10가지 징후에 초점을 맞춥니다 [14]. 참가자는 0에서 6가지의 점수로 질문에 답합니다. 0~6점은 정상 또는 증상 없음, 7~19점은 경증, 20~34점은 중등도, 35~60점은 중증 우울증으로 분류됩니다 [15].

치료 반응에서도 개인차가 나타납니다. 어떤 사람은 치료에 잘 반응하는 반면, 같은 증상을 가진 다른 사람은 같은 치료에 같은 방식으로 반응하지 않을 수 있습니다. 따라서 진단과 치료가 지연될 경우 심각한 합병증을 유발하여 개인의 생명을 위협할 수 있습니다 [16]. 우울증 진단은 매우 어렵고, 다른 유형의 질환과 유사한 증상을 보이는 경우가 많아 오진되는 경우가 흔합니다 [17]. 예를 들어 불안 장애로 오진되는 경우가 많습니다 [18]. 그러므로 우울증의 조기 발견 및 진단이 매우 중요합니다.

지난 몇 년 동안 인공지능(AI) 분야는 발전해 왔으며, 의료 분야를 포함한 다양한 분야에서 성공을 거두었습니다 [19]. 머신러닝(ML, AI의 하위 분야)은 개인이 정신 질환을 앓고 있는지 여부와 자살 경향성을 예측할 수 있는 잠재력을 가지고 있습니다 [20]. 따라서 ML을 구현하면 취약한 사람들을 식별하여 우울증을 포함한 정신 건강 장애를 조기에 발견하는 데 도움이 되어 정신 질환 분야에 유익할 수 있습니다.

기존의 머신러닝 기반 우울증 탐지 모델은 다양한 유형의 우울증을 정확하게 구분하지 못하는 등 여러 가지 문제점을 안고 있습니다. 심층 신경망과 같은 더욱 발전되고 정교한 기술조차도 "블랙박스" 접근 방식으로 인해 투명성과 설명 가능성이 부족하여 사용에 제약이 따르는 경우가 많습니다 [19]. 이러한 해석 가능성의 부족은 의사 결정 과정을 이해하는 것이 매우 중요한 의료 분야에서 큰 걸림돌이 됩니다 [19]. 특히 우울증의 진단 및 치료 맥락에서는 투명성과 설명 가능성이 더욱 중요합니다.

우울증에 대한 AI 모델은 효과적인 모델링을 위해 행동적, 심리적, 생리적 지표를 포착하는 방대한 양의 데이터를 필요로 하며, 따라서 우울증 상태를 정확하게 감지하고 구분할 수 있습니다. 그러나 이러한 데이터 세트의 수집은 데이터 수집 방법의 어려움과 한계로 인해 종종 불가능합니다 [21].

대안은 신체 활동, 수면 패턴, 심박 변이도 및 기타 여러 생리적 지표와 같은 다양한 측면에 대한 지속적인 실시간 데이터를 수집하기 위해 웨어러블 장치를 사용하는 것입니다 [21].

다양한 데이터셋과 머신러닝 모델을 사용하여 여러 유형의 정신 질환 진단에 대한 연구가 다수 진행되었습니다. Paul과 Juliet [22]은 환자들이 직접 보고한 정보 데이터셋을 기반으로 정신 건강 장애 문제를 예측하는 것을 고려했습니다. 저자들은 로지스틱 회귀, k-최근접 이웃, 의사결정 트리, 랜덤 포레스트(RF)의 네 가지 머신러닝 알고리즘을 사용했습니다. RF는 81%의 정확도를 달성하여 가장 우수한 성능을 보이는 알고리즘으로 확인되었습니다. 그러나 저자들은 재현율, 정밀도, F1 점수와 같은 다른 성능 지표에 대해서는 논의하지 않았습니다.

Zhang et al. [23]은 외상 후 스트레스 장애를 분류하기 위해 자기파착정(MEG)을 사용하는 아키텍처를 제안했습니다. 저자들은 재귀적 랜덤 포레스트(RF)를 특징 선택에 사용하는 서포트 벡터 머신(SVM)을 주요 분류기로 사용했습니다. 특징 선택이 이루어진 모델은 AUC(곡선 아래 면적) 값이 최대 0.9에 도달했습니다. Mikolas et al. [24]은 주의력 결핍 과잉 행동 장애(ADHD)의 자동 진단에 초점을 맞추었으며, 머신러닝 분야의 최근 발전이 ADHD와 정상인을 구분하는 이진 분류 문제에는 도움이 되었지만, 여러 정신 질환을 구별하는 문제는 여전히 남아 있다고 주장했습니다. 저자들은 익명화된 참가자 229명의 데이터셋을 고려하여 자동 특징 선택을 사용할 때 최상의 성능을 얻을 수 있음을 관찰했습니다.

본 연구는 특정 열에 결측값이 많고 표본 외 검증이 부족하다는 등의 몇 가지 한계점을 가지고 있습니다.

Spulber et al [25]은 활동 기록계를 통해 수집된 활동 데이터와 우울증 심각도 수준이 상관관계가 있는지 조사했습니다.

이 연구는 항우울제 치료를 받지 않고 있는 주요 우울 삽화를 경험하는 사람들을 대상으로 진행되었습니다.

저자들은 머신러닝 기법을 사용하여 활동 데이터로부터 MADRS 점수를 예측하는 모델을 훈련시키고 정확도와 정밀도를 기준으로 모델을 선택했습니다. 이러한 모델은 평균 제곱근 오차(RMSE) 값을 기준으로 최상, 중간, 최악으로 순위를 매겼습니다. 치료를 받고 있는 환자의 데이터를 사용하여 외부 검증도 수행했습니다. 그러나 이 외부 검증은 환자 상태 및 치료 상태의 차이로 인해 잠재적인 편향을 초래할 수 있습니다. Garcia-Ceja 등 [26]은 대조군과 치료군의 활동량 측정 데이터가 포함된 Depresjon 데이터셋을 사용하여 정신운동 데이터로부터 건강한 참가자와 건강하지 않은 참가자를 식별하기 위해 랜덤 포레스트(RF)와 심층 신경망(DNN)의 두 모델을 비교했습니다. 클래스 불균형으로 인해 랜덤 오버샘플링 및 합성 소수 클래스 오버샘플링 기법(SMOTE)과 같은 다양한 샘플링 방법을 채택하고 세 가지 통계적 특징을 추출한 후 기준선 및 오버샘플링을 하지 않은 경우와 비교했습니다. 그 결과 SMOTE를 적용한 RF 모델이 0.73의 더 높은 F1 점수를 달성한 것으로 나타났습니다.

Frogner et al [27]은 활동 데이터를 사용하여 우울증을 감지하기 위해 1D 컨볼루션 신경망(CNN)을 구현했습니다.

모델은 서로 다른 시간 구간으로 학습되었으며, 48시간 구간이 최적의 것으로 나타나 3가지 분류 작업을 수행했습니다. 첫 번째 작업은 우울증 그룹을 분류하고, 두 번째 작업은 우울증 수준(MADRS 점수 범위 0-10, 11-19, >20을 기준으로 우울증 없음, 경도 우울증, 중등도 우울증)을 분류하며, 마지막 작업은 MADRS 점수를 예측합니다. 평가는 leave-one-participant-out 교차 검증과 다수결 투표를 결합하여 수행되었습니다. 첫 번째 작업에서 F1 점수 70%, 두 번째 작업에서 30%, 마지막 작업에서 평균 제곱 오차(MSE) 4.0을 달성했습니다. 모델은 우울증이 없는 참가자를 분류하는 데는 우수한 성능을 보였지만, 경도 및 중등도 우울증 수준을 분류하는 데는 어려움을 겪었습니다. Jakobsen et al. [28]은 ML 모델을 사용하여 활동 패턴을 분석함으로써 활동 데이터의 객관적 지표가 현재의 우울증 진단 기법을 향상시킬 수 있는지 여부를 조사했습니다. RF 및 DNN 모델에 대해 세 가지 통계적 속성(평균 활동 수준, 표준 편차, 활동이 없는 비율)을 도출했습니다. CNN은 24×60 이미지 행렬로 표현된 원시 활동 데이터를 사용하여 학습되었습니다. SMOTE를 적용한 DNN은 84%의 정확도, 82%의 참양성률, 84%의 참음성률을 달성했습니다. 그러나 특히 초조 증상을 보이는 환자와 같이 활동 패턴이 중복되는 그룹에서 상당한 오분류가 관찰되었습니다.

Espino-Salinas 등 [29]은 Depresjon 데이터셋을 사용하여 우울증 분류 프레임워크를 개발했습니다. 전처리 과정에서는 운동 활동 데이터를 벡터로 변환하고 신경망 입력용 행렬을 생성했습니다. 이후 2D-CNN과 DNN을 사용하여 우울증 환자와 비우울증 환자를 분류했습니다. 2D-CNN은 76%의 정확도와 72%의 F1 점수를 달성했고, DNN은 72%의 정확도와 69%의 F1 점수를 달성했습니다. 데이터셋의 불균형을 고려하여 F1 점수를 주요 평가 지표로 선택했는데, 이는 모델이 불균등한 클래스 분포를 처리하는 능력을 더 잘 반영하기 때문입니다.

본포. 2D-CNN은 민감도와 특이도 사이의 더 나은 균형을 달성하여 DNN보다 우수한 성능을 보였으며, 우울증 분류에 대한 잠재력을 입증했습니다. 지면 제약을 고려하여 자세한 내용은 [30-35]를 참조하십시오.

정신 건강 진단을 위한 머신러닝(ML) 기술이 상당한 발전을 이루었음에도 불구하고, 여전히 몇 가지 한계점이 남아 있습니다. 여러 연구에서 반복적으로 나타나는 한계는, 특히 우울증 심각도 수준과 같이 중복되거나 미묘한 차이를 보이는 질환을 구분하는 데 있어 상대적으로 낮은 성능입니다. 모델은 이진 분류(예: 우울증 환자와 비우울증 환자 구분)에서는 우수한 성능을 보이지만, 다중 클래스 분류나 경증 우울증과 중등도 우울증처럼 질환의 미묘한 차이를 구분하는 데 어려움을 겪습니다. 또 다른 중요한 한계는 이러한 연구에서 설명 가능한 인공지능(XAI) 기술이 부족하다는 점입니다. CNN과 DNN을 포함한 현재의 ML 모델은 블랙박스 방식으로 작동합니다. 블랙박스 방식은 이러한 기술이 의사 결정 과정을 제대로 이해하지 못하게 만듭니다. 이러한 해석 불가능성은 신뢰도를 떨어뜨리고, 모델이 특정 예측을 하는 이유를 이해하는 것이 예측 자체만큼 중요한 임상 환경에서의 실제 적용을 제한합니다.

본 연구에서는 웨어러블 기기에서 수집한 데이터를 활용하여 우울증 탐지 및 심각도 분류 작업에 머신러닝 알고리즘을 적용하는 방안을 조사합니다.

불균형 데이터셋 문제를 해결하기 위해 적응형 합성 샘플링 기법(ADASYN)을 사용했습니다. 또한, 활동 데이터에서 의미 있는 특징을 추출하기 위해 특징 엔지니어링을 수행했습니다. 이진 분류(우울증: 있음 또는 없음)와 다중 클래스 분류(우울증 심각도 예측)를 위해 로지스틱 회귀, 서포트 벡터 머신, 랜덤 포레스트(RF), XGBoost 및 신경망을 구현했습니다. 모델 성능은 정확도, 정밀도, 재현율, F1 점수 및 매튜 상관계수(MCC)를 사용하여 평가했습니다. 모델의 설명력을 더욱 향상시키기 위해, 가장 우수한 성능을 보인 모델들의 의사결정 과정을 설명하기 위해 샵러리 가산 설명(SHAP)과 지역 해석 가능 모델 불가지론적 설명(LIME) 방법을 사용했습니다.

행동 양식

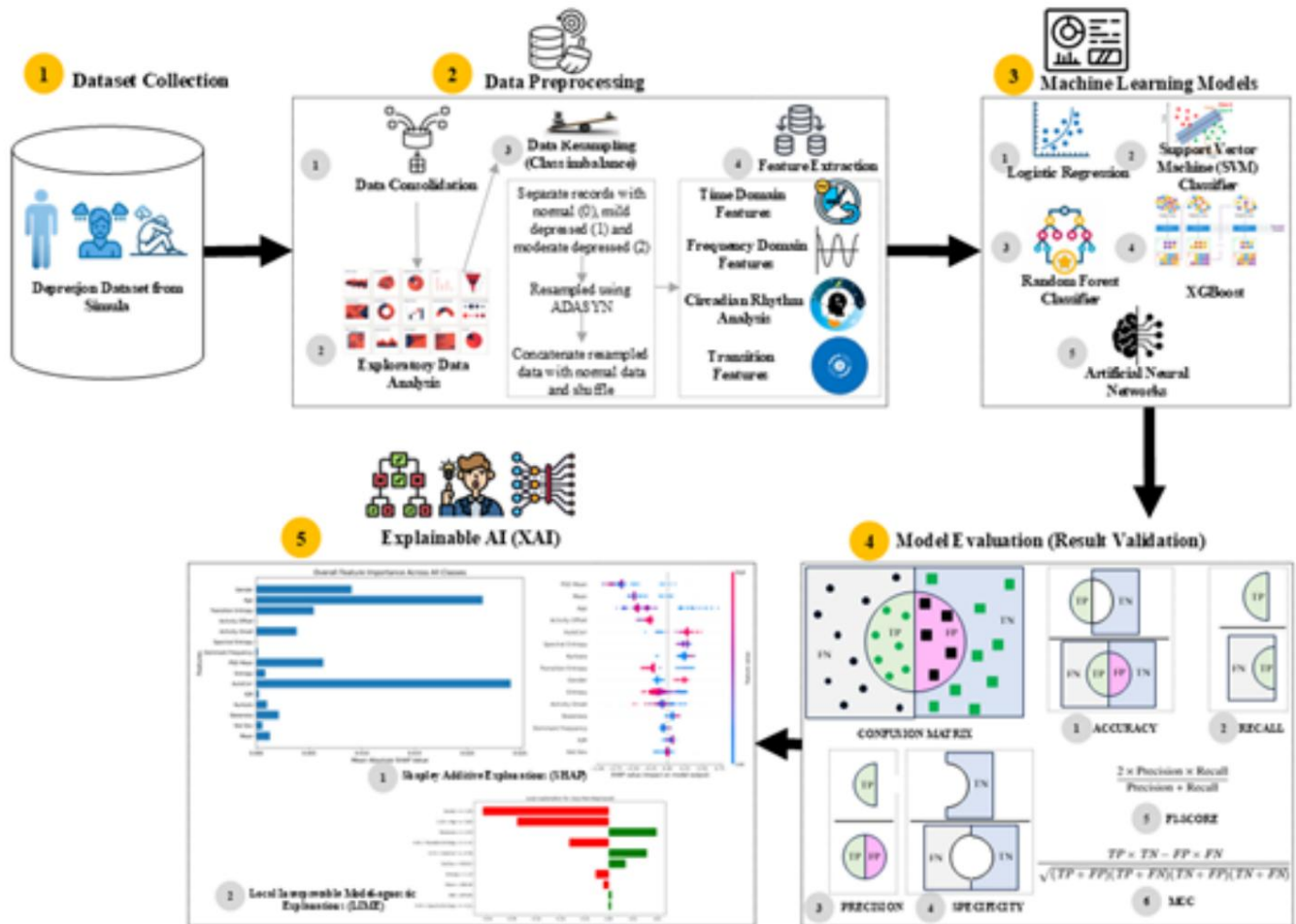
개요

그림 1은 본 연구 방법론을 도식적으로 나타낸 것입니다.

우리는 Depresjon 데이터셋에서 데이터를 확보하고 탐색적 데이터 분석을 수행하여 데이터의 특성을 이해하고자 했습니다.

데이터에서 클래스 불균형이 관찰되었으며, 샘플링 기법을 사용하여 이를 해결한 후 데이터 정규화를 수행했습니다. 그런 다음 특징 엔지니어링을 적용하여 데이터에서 특징을 추출했습니다. 이어서 이진 분류(우울증 vs. 비우울증) 및 다중 클래스 분류(우울증 심각도 수준) 작업을 위한 모델을 설계했습니다. 모델 구현 후 정확도, 정밀도, 재현율, 특이도, MCC 및 F1 점수를 성능 평가 기준으로 사용했습니다. 마지막으로 SHAP 및 LIME 방법을 적용하여 선택된 모델의 의사 결정 과정을 설명했습니다.

그림 1. 방법론의 도식적 표현.



데이터 수집

본 연구에서 사용된 데이터셋은 Simula 저장소 [36]에서 이용 가능한 "Depresjon"(노르웨이어로 우울증) 데이터셋입니다. 이 데이터셋은 활동량 측정 시계를 사용하여 며칠 동안 분 단위로 기록된 개인의 운동 활동 데이터로 구성됩니다. 데이터셋은 두 부분으로 나뉩니다. 첫 번째 부분은 조건군과 대조군의 각 개인에 대한 활동 수준 측정값으로 이루어져 있으며, 파일은 타임스탬프, 날짜, 활동량 열로 구성됩니다. 두 번째 부분은 각 참가자의 인구통계학적 정보를 포함합니다. 데이터셋에는 식별 번호(number), 측정 일수(days), 성별(여성 1, 남성 2), 참가자의 연령대(age)와 같은 여러 변수가 포함되어 있습니다. 또한 정서 장애 유형(aftype)도 기록되어 있습니다. 멜랑콜리아 증상은 1(멜랑콜리아 있음)과 2(멜랑콜리아 없음)로 표시됩니다. 데이터셋은 입원 환자를 '1', 외래 환자를 '2'(inpatient)로 구분합니다. 교육 수준은 이수 연수로 분류됩니다(uceducation). 혼인 여부는 기혼 또는 동거인 경우 1, 미혼(결혼)인 경우 2로 기록됩니다.

일관된 데이터셋입니다. 병합 작업 후 생성된 데이터셋은 총 1440개의 데이터 포인트로 구성됩니다. 이진 분류(우울증 vs 비우울증)의 경우, "조건군"은 1, "대조군"은 0으로 레이블링되었습니다. 다중 클래스 분류(우울증 심각도 평가)의 경우, 정상은 0, 경도 우울증은 1, 중등도 우울증은 2로 레이블링되었습니다. 참가자는 MADRS 점수에 따라 0~6점은 "정상"(우울증 없음), 7~19점은 "경도 우울증", 20~34점은 "중등도 우울증"으로 분류되었습니다.

표준 MADRS 척도에는 35점 이상에 대한 "심각" 범주가 포함되어 있지만, 저희 데이터 세트에는 심 각 범위에 해당하는 점수가 없었으므로 "심각" 범주를 포함하지 않았습니다.

데이터 불균형 문제 해결

우리는 데이터가 불균형하다는 것을 관찰했는데, 이는 일부 출력 클래스가 다른 클래스보다 더 많이 나타난다는 것을 의미합니다. 불균형 데이터는 ML 알고리즘의 성능을 왜곡할 수 있으므로 해결이 필요합니다. 우리는 클래스 불균형 문제를 해결하기 위해 ADASYN을 사용했습니다. ADASYN의 기본 아이디어는 학습 난이도에 따라 빈도가 낮은 인스턴스에 가중치를 부여하는 것입니다 [37]. ADASYN의 작동 방식에 대해서는 He et al. [37]을 참조하십시오.

근무 상태는 취업 또는 학업 중인 경우 1, 실업 또는 병가 또는 연금 수급(취업) 중인 경우 2로 코딩됩니다. 마지막으로, 데이터 세트에는 검사 전(madrs1)과 후(madrs2)의 MADRS 점수가 포함됩니다.

원래 데이터 세트는 두 개의 폴더(조건군과 대조군)로 구성되어 있었으며, 데이터를 하나의 폴더로 병합하는 것이 중요했습니다.

특징 엔지니어링

이 데이터 세트는 활동량 측정 시계로 기록된 활동 측정값으로 구성되어 있으며, 각 환자에 대해 일정한 간격으로 수집된 연속적인 데이터 스트림을 제공합니다. 그러나 이러한 원시 시계열 값을 사용하는 것은 불필요하게 많은 상호 연결된 특징들을 생성하고 정신 건강 상태를 나타내는 미묘한 행동 및 생리적 리듬을 포착하지 못하기 때문에 효과적인 접근 방식이 아닙니다.

우울증 및 관련 기분 장애의 맥락에서 일상 활동 패턴의 미묘한 변화, 주파수에 걸친 에너지 분포, 비선형 변동은 특히 중요한 의미를 지닙니다. 이를 해결하기 위해 연속 시계열 데이터를 주요 통계적 특징으로 요약하여 원시 센서 데이터를 심리적 및 임상적 통찰력과 더 밀접하게 연관된 지표로 변환했습니다. 이러한 특징 엔지니어링 접근 방식은 필수적인 패턴과 변동성을 유지하면서 연속 데이터 스트림을 관리하는 표준 메커니즘 [38]으로, 데이터 차원과 계산 요구 사항을 효과적으로 줄입니다.

인간 활동의 다면적인 특성과 우울 증상과의 잠재적 연관성을 포착하기 위해, 우리는 5가지 광범위한 범주에서 특징을 체계적으로 추출했습니다: (1) 시간 영역, (2) 주파수 영역, (3) 일주기 리듬 분석, (4) 전환 특징. 각 범주는 정신 건강 상태와 흔히 연관되는 기분 변화나 일상 활동의 중단을 반영할 수 있는 활동의 다양한 측면을 다룹니다.

시간 영역 특징

시간 영역 특징은 활동 신호의 통계적 속성과 변동성을 설명합니다. 이러한 측정값의 교란 또는 변동성 증가는 기분 장애를 겪는 개인에게서 흔히 관찰되는 에너지 수준 및 일상생활의 변화와 연관될 수 있습니다. 시간 영역 특징(평균 활동, 변동성, 고차 모멘트, 자기상관 및 엔트로피 등)은 오랫동안 활동량 측정 연구의 핵심 요소였습니다. 기존 연구에 따르면 이러한 측정값은 우울증과 비우울증의 운동 패턴을 일관되게 구분하여 정신운동 변화에 대한 간결하고 행동적으로 해석 가능한 정보를 제공합니다 [25-27,35,38].

다음은 시간 영역 특징들을 추출한 결과입니다.
지속적인 데이터 스트림.

- 평균 \pm 표준편차: 기록된 활동 수준의 중심 경향과 전반적인 변동성을 요약합니다. 현저한 변동은 우울증 삽화에서 흔히 나타나는 불규칙한 수면-각성 주기 또는 에너지 불안정을 나타낼 수 있습니다.
- 왜도와 첨도: 활동 분포의 비대칭성과 "뾰족함"을 나타냅니다. 이러한 지표의 급격한 변화는 무기력이나 과잉행동기를 나타낼 수 있으며, 이는 기분 상태의 변화를 반영할 수 있습니다.
- IQR: 25번째 백분위수와 25번째 백분위수 사이의 분포를 측정합니다. 75번째 백분위수는 일상 활동이 얼마나 밀집되어 있는지 또는 분산되어 있는지를 명확하게 보여주는 지표이며, 이는 기분 변화와 상관관계가 있을 수 있습니다. • 자기상관(AutoCorr): 유사성을 분석합니다.

활동 수준은 다양한 시간 지연을 두고 자체적으로 나타납니다.

- 반복적인 패턴이나 정신 건강 장애에서 흔히 나타나는 일주기 리듬 장애를 감지하는 데 도움이 됩니다. • 엔트로피: 활동 패턴의 무작위성 또는 불

규칙성을 평가합니다. 엔트로피가 높다는 것은 불규칙적인 일상 스케줄을 시사하며, 이는 수면 위생 불량이나 기분 불안정의 증상일 수 있습니다.

주파수 영역 특징

다양한 주파수에 걸쳐 활동 전력 분포를 분석하면 기저 주기성을 밝힐 수 있습니다. 이러한 주기적 구성 요소의 변화는 종종 수면-각성 주기 장애 또는 비정상적인 에너지 분포와 관련이 있으며 이는 우울증에서 흔히 나타납니다. 여러 활동량 측정 연구에서 이러한 스펙트럼 이상이 증상 심각도와 연관되어 있으며 기분 추론에 대한 진단적 가치가 확인되었습니다 [25,29,35,38].

추출된 특징은 다음과 같습니다. • 전력

스펙트럼 밀도(평균): 신호의 전력이 주파수에 걸쳐 어떻게 분포되는지를 나타내는 척도로, 일상 활동의 주기적 구성 요소의 전반적인 강도를 반영합니다 [39]. • 지배 주파수: 활동이 가장 두드러지게 나타나는 주파수를 식별하며, 일반적으로 일주기(생체 리듬)와 관련이 있습니다. 편차는 우울증 증상이 있는 개인의 일주기 리듬이 교란되거나 변화되었음을 반영할 수 있습니다.

- 스펙트럼 엔트로피: "얼마나 퍼져 있는지" 또는 활동력은 여러 주파수 대역에 걸쳐 집중되어 있습니다. 확산 정도가 커지면 일상생활이 불안정해지고 패턴이 더 불규칙해질 수 있으며, 이는 기분 변화와 관련될 가능성이 있습니다.

일주기 리듬 분석

기분 장애는 종종 일주기 리듬의 교란을 동반합니다. 주요 전환 시기를 조사하면 수면-각성 불규칙성에 대한 귀중한 단서를 얻을 수 있습니다. [40]

(활동 시작 및 종료: 주요 일상 활동 기간의 시작과 끝을 파악합니다. 이 값의 변화는 수면 단계의 지연 또는 앞당김을 나타낼 수 있으며, 이는 우울증이나 기타 정서 장애가 있는 사람들에게서 흔히 관찰됩니다.)

전환 기능

활동은 하루 종일 항상 매끄럽게 변화하는 것은 아닙니다. 갑작스러운 전환이나 안정적인 패턴은 특히 정신 건강 상태를 잘 나타낼 수 있습니다. 전환 엔트로피는 활동 상태 전환 순서에 내재된 불확실성 또는 예측 불가능성을 포착합니다 [41]. 전환 엔트로피는 전환 확률 분포를 평가하여 서로 다른 활동 상태 간의 전환에서 무작위성을 정량화합니다.

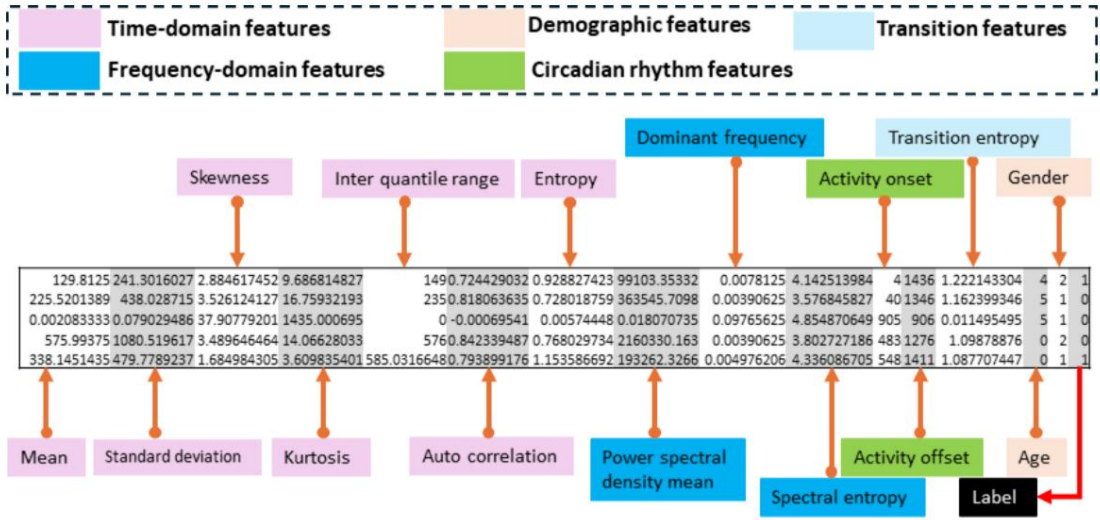
높은 전환 엔트로피는 일상생활의 불안정성이나 기분 변화를 반영할 수 있으며, 이는 특정 정신 질환을 가진 사람들에게서 흔히 관찰됩니다.

시간 영역 및 주파수 영역 측정 지표와 더불어, 인간 활동 신호에서 흔히 관찰되는 자기 유사 구조와 장기 의존성을 더 잘 포착하기 위해 비선형 기술 지표(특히 프랙탈 차원 및 허스트 지수) 세트를 테스트했습니다. 이를 위해 여러 가지 방법을 구현했습니다.

공개된 알고리즘을 사용하고 여러 윈도우 길이에 걸쳐 매개변수를 조정했습니다. 이러한 노력에도 불구하고 추출 파이피라인은 지속적으로 정의되지 않거나 0인 값을 많이 생성하여 데이터 해석도의 수치적 불안정성을 나타냈습니다. 문제 해결 및 민감도 검사 후에도 문제는 지속되었으며 모델의 신뢰성을 위협했습니다.

견고성과 해석 가능성을 확보하기 위해, 본 연구에서는 비선형 분석을 시도한 부분을 기록하고 최종 특징 집합에서 해당 특징들을 제외했습니다. **그림 2** 는 특징 엔지니어링 후의 데이터셋을 도식적으로 나타낸 것입니다.

그림 2. 특징 추출 후 데이터의 스냅샷.



특징 엔지니어링 후, 우리는 식 (1) 에 나타낸 바와 같이 z-점수 정규화 기법을 사용하여 데이터 정규화를 수행했습니다 .

=
$$\frac{x_i - \mu}{\sigma}$$

여기서 xi 와 zi는 표준화 전후의 개별 특성 점수를 나타내고, μ와 σ는 활동 수준 점수의 평균과 표준 편차를 나타냅니다.

로지스틱 회귀

로지스틱 회귀(LR)는 주로 분류 작업에 사용되는 널리 사용되는 지도 학습 알고리즘입니다 [42].

로지스틱 회귀 분석(LR)은 입력 변수가 특정 클래스에 속할 확률을 추정하기 위해 입력 변수 특징들의 선형 조합에 로지스틱(시그모이드) 함수를 적용합니다. LR은 특히 특징 변수와 목표 변수 간의 관계가 대략 선형일 때 단순성, 해석 용이성, 효율성 측면에서 선호됩니다. 또한 확률적 통찰력을 제공하여 예측의 신뢰도를 평가할 수 있도록 합니다.

또한 LR은 One-vs-Rest(OvR) 또는 다항 LR [43]과 같은 전략을 사용하여 다중 클래스 분류 문제로 확장될 수 있습니다. 단순함에도 불구하고 LR은 강력한 기준 모델 역할을 하며 결정 경계가 선형이거나 거의 선형인 시나리오에서 효과적입니다.

서포트 벡터 머신

서포트 벡터 머신(SVM)은 특히 분류 문제에 널리 사용되는 모델입니다. SVM의 핵심 개념은 n차원 공간에서 데이터 포인트를 서로 다른 클래스로 분리하는 최적의 초평면을 찾는 것입니다 [44]. 이진 분류 문제의 경우, SVM은 두 클래스를 분리하는 최적의 초평면을 찾습니다.

특징 공간. 다중 클래스 문제의 경우 일반적으로 OvR과 One-vs-One(OvO)의 두 가지 접근 방식이 사용됩니다. OvR에서는 각 클래스에 대해 다른 모든 클래스와 이진 분류기를 사용하는 반면, OvO에서는 가능한 모든 클래스 쌍에 대해 이진 분류기를 학습합니다 [44]. 즉, k개 클래스 분류 문제의 경우 OvO는 이진 분류기를 학습합니다.

$$\frac{N(N-1)}{2}$$

랜덤 포레스트

RF는 여러 의사결정 트리의 결과를 통합하여 보다 정확한 결과를 도출하는 지도 앙상블 학습 알고리즘입니다 [45]. 의사결정 트리와 달리 랜덤 포레스트는 과적합을 최소화하는 능력을 가지고 있습니다 [45]. RF는 배깅(bagging) 방식을 기반으로 하며, 무작위로 선택된 특징 부분집합을 사용하여 여러 모델을 학습시키고, 결과 예측값을 투표 또는 평균 메커니즘을 통해 집계합니다. 회귀의 경우 예측값은 모든 의사결정 트리의 평균 예측값이며, 분류의 경우 의사결정 트리에서 다수 결로 예측된 클래스입니다 [46].

XGBoost

XGBoost(Extreme gradient boosting)는 Gradient Boosting 프레임워크 [47] 를 기반으로 하는 ML 알고리즘입니다 . Gradient Boosting의 아이디어는 최종적으로 강력한 학습 모델을 얻기 위해 결과를 집계할 수 있을 것으로 기대되는 여러 개의 약한 학습기(의사결정 트리와 같은)를 개발하는 데 기반을 두고 있습니다 [47].

XGBoost는 의사 결정 트리의 앙상블을 순차적으로 구축하며, 각 새로운 트리는 이전 트리의 오류를 수정하려고 시도합니다. XGBoost는 특히 불균형 데이터 세트를 처리하는 데 유용합니다 [48].

신경망

신경망(NN)은 인간 두뇌의 뉴런 네트워크에서 영감을 받은 강력하고 유연한 머신러닝 모델의 한 종류입니다. 상호 연결된 노드(뉴런)의 계층으로 구성된 신경망은 입력 특징과 목표 변수 간의 복잡하고 비선형적인 관계를 모델링할 수 있습니다.

각 뉴런은 입력의 가중합에 이어 비선형 활성화 함수를 적용하여 네트워크가 데이터로부터 복잡한 패턴과 표현을 학습할 수 있도록 합니다 [49].

정신 건강 분석의 맥락에서 신경망은 활동량 측정 데이터에 존재하는 복잡한 시간적, 공간적 종속성을 효과적으로 포착할 수 있으므로 우울증 분류 및 심각도 평가와 같은 작업에 매우 적합합니다 [50]. 원시 데이터에서 특징 표현을 자동으로 학습하는 능력 덕분에 광범위한 특징 엔지니어링의 필요성이 줄어들지만, 일반적으로는

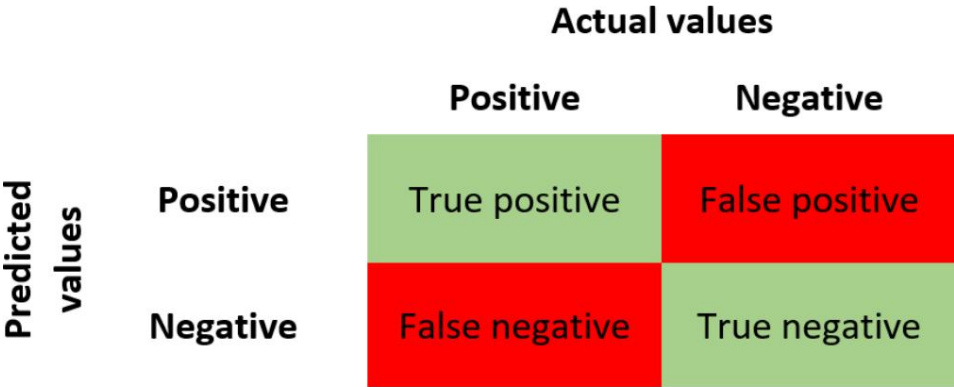
기존 알고리즘에 비해 더 큰 데이터 세트와 더 많은 컴퓨팅 리소스를 필요로 합니다. 하이퍼파라미터 설정은 [멀티미디어 부록 1](#)에 보고되어 있습니다.

하이퍼파라미터 설정은 [멀티미디어 부록 1](#)에 보고되어 있습니다.

평가 지표

본 연구에서는 ML 모델의 성능 평가 기준으로 정확도, 재현율, 특이도, 정밀도, F1-점수 및 MCC를 사용합니다. 이러한 성능 평가 지표는 분류 문제에 대한 정확하게 분류된 인스턴스와 잘못 분류된 인스턴스를 요약한 혼동 행렬을 기반으로 합니다. 혼동 행렬의 그래픽 표현은 [그림 3](#)에 나와 있습니다. 혼동 행렬은 참양성(TP), 거짓양성(FP), 거짓음성(FN) 및 참음성(TN)이라는 4가지 핵심 요소로 구성되며 다음과 같이 정의됩니다 [51].

그림 3. 혼동 행렬.



TP는 모델이 긍정 사례를 올바르게 분류한 경우입니다. FP는 모델이 부정 사례를 긍정 사례로 분류한 경우입니다. FN은 모델이 긍정 사례를 부정 사례로 분류한 경우입니다.

TN은 모델이 부정 인스턴스를 부정 인스턴스로 분류하는 경우를 나타냅니다. 혼동 행렬의 다양한 요소에 대한 정의를 바탕으로 [표 1](#)에서 성능 평가 지표를 정의합니다.

표 1. 성능 지표.

미터법	공식
정확도(Accc)	$\frac{TP + TN}{TP + TN + FP + FN}$
리콜(Rec)	$\frac{TP}{TP + FN}$
특이성(Spec)	$\frac{TN}{TN + FP}$
정밀도(Prec)	$\frac{TP}{TP + FP}$
F1 점수(F1)	$2 \times \frac{\text{정밀도} \times \text{재현율}}{\text{정밀도} + \text{재현율}}$
MCC	$\frac{(TP \times TN) - (FP \times FN)}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}$

결과

바이너리 및 성능 평가 다중 클래스 분류 작업

표 2는 이진 분류 작업에서 다양한 머신러닝 모델의 성능을 보여줍니다. 평가된 모델들 중에서 XGBoost가 84.94%로 가장 높은 정확도를 달성했습니다. 다른 모든 모델보다 훨씬 뛰어난 성능을 보였습니다. 이와 대조적으로 LR은 가장 효과가 떨어졌으며, 정확도는 70.68%였습니다.

표 2. 다양한 모델의 성능 비교.

모델	정확성	정도	상기하다	F1 점수	매튜 상관관계 계수
로지스틱 회귀	0.7068	0.6937	0.7471	0.7187	0.4156
SVM	0.7789	0.7648	0.8056	0.7842	0.5595
XGBoost	0.8494	0.8340	0.8745	0.8529	0.7012
랜덤 포레스트	0.8301	0.8096	0.8663	0.8358	0.6639
신경망	0.8368	0.8275	0.8543	0.8397	0.6758

또한 XGBoost는 기존 제품 대비 향상된 성능을 보였습니다. 다른 모델들의 평균 성능을 모든 평가 지표에서 뛰어넘었습니다. 특히 XGBoost는 평균을 능가하는 성능을 보였습니다. 정확도, 정밀도, 재현율, F1 점수 및 MCC가 7.2% 향상되었습니다. 각각 6.42%, 6.83%, 17.47% 향상되었습니다. 이러한 개선은 XGBoost가 신호 포착에 있어 탁월한 성능을 발휘함을 보여줍니다. 데이터 내의 복잡한 패턴과 상호작용으로 인해 본 연구에서 가장 효과적인 모델입니다. 이러한 결과는 XGBoost와 같은 앙상블 방법의 효과를 입증합니다. 기존 모델 및 기타 머신러닝 모델보다 우수한 성능을 보임으로써, 향상된 예측 정확도와 신뢰성을 제공합니다.

표 3은 다양한 머신러닝 모델의 성능을 보여줍니다. 다중 클래스 분류 작업에서 XGBoost는 다음과 같은 성과를 달성했습니다.

표 3. 다중 클래스 분류를 위한 다양한 모델의 성능 비교.

모델	정확성	정도	상기하다	F1	매튜 상관관계 끊임없는
로지스틱 회귀	0.5412	0.5414	0.5412	0.5386	0.3136
SVM	0.5322	0.5362	0.5322	0.5315	0.2999
XGBoost	0.8591	0.8602	0.8591	0.8581	0.7895
랜덤 포레스트	0.8285	0.8308	0.8285	0.8276	0.7446
신경망	0.8352	0.8396	0.8352	0.8342	0.7554

통계적 유의성 검증

XGBoost의 우수한 성능을 검증하기 위해 이는 우연의 일치가 아니라, 아래에 자세히 설명된 바와 같이 엄격한 통계적 유의성 검정을 수행한 결과입니다. 10×3 표본 추출을 실시했습니다. 반복적인 총화 교차 검증 프로토콜(총 30검) 분할할 때마다 원래의 클래스 분포를 유지합니다. 각 폴드에서 모델은 훈련 지표에 맞춰 조정되었으며, 검증 데이터셋에 대한 이진 예측은 다음과 같이 점수가 매겨졌습니다. F1 지표는 적당히 불균형한 경우에 적합합니다. 문제점을 파악하기 위해, 우리는 쌍측 검정 t 검정을 적용했습니다. XGB와 각 경쟁 모델, 그리고 홀름 보정을 차례로 적용했습니다. 전체 오류율을 $\alpha=0.05$ 로 제한합니다.

정확도를 기준으로 한 순위는 모든 항목에 일관되게 반영됩니다. 기타 지표: XGBoost는 정확도(83.40%)에서도 선두를 차지합니다. 회상률(87.45%), F1 점수(85.29%), MCC(0.7012) 반면 LR은 이러한 모든 범주에서 가장 낮은 점수를 받았습니다. SVM, RF 및 신경망을 포함한 나머지 모델 네트워크는 성능이 점진적으로 향상되는 것을 보여줍니다. SVM이 로지스틱 회귀(LR), 랜덤 포레스트(RF) 및 신경망보다 우수한 성능을 보였습니다. XGBoost를 바짝 뒤쫓고 있습니다. 이러한 모든 분야에서의 일관된 추세는 평가 지표는 XGBoost의 견고성을 강조합니다. 이 분류 작업에 가장 적합한 모델입니다.

정밀도, 재현율과 함께 최고 정확도 85.91%를 기록했습니다. F1 점수가 85%를 초과 하고 MCC가 0.7895인 경우. 반면 SVM과 LR은 가장 낮은 정확도를 기록했습니다. 각각 53.22%와 54.12%를 기록하며 저조한 성과를 보였습니다. 다른 지표에서도 마찬가지입니다. RF 및 신경망 82.85%의 정확도로 우수한 성능을 보여주었습니다. 각각 83.52%와 83.52%로, 일관되게 더 높은 정밀도를 보였습니다. 회상률, F1 점수 및 MCC. 이러한 결과는 다음을 나타냅니다. XGBoost는 평균 정확도, 정밀도, 재현율을 뛰어넘었습니다. F1 점수 및 MCC는 각각 20.35%, 20.13%, 20.34%, 20.41% 증가했습니다. 그리고 각각 33.07%입니다.

이진 분류의 경우, 30번의 교차 검증을 거쳤습니다. 접힘 수에서 XGB는 가장 높은 평균 F1 점수(0.851 ± 0.034)인 반면 가장 가까운 경쟁자인 RF는 평균 0.832를 기록했습니다. SVM, LR과 NN은 각각 0.659, 0.723, 0.687의 수치를 기록했습니다. 쌍별 유의성 검정을 통해 이러한 차이가 확인되었습니다. 샘플링 노이즈 때문이 아닙니다: Holm으로 조정된 P 값 XGB와 RF, LR, SVM, NN의 비교에서 R^2 값은 3.6×10^{-2} 였습니다. 2.3×10^{-1} , 2.7×10^{-1} , 및 1.7×10^{-2} 는 모두 0.05보다 훨씬 낮습니다. XGBoost의 우수한 성능을 확인시켜 줍니다. 우연이 아닙니다. 따라서 XGB의 성능 순위는 다음과 같습니다. 다른 모든 기준선과 비교했을 때 통계적으로 유의미함 연구 결과는 해당 모델이 주요 모델로 선정되었음을 뒷받침합니다. 후속 설명 가능성 분석. 동일한 추세가 나타났습니다. 다중 클래스 분류에 대해 관찰되었습니다.

논의

주요 결과

이 섹션에서는 XAI 기법(SHAP 및 LIME)을 비판적으로 해석하여 우울증 분류 모델을 구동하는 특징 패턴을 밝히고, XGBoost와 신경망 아키텍처 전반에 걸친 이러한 설명들을 비교하며, 이러한 결과들을 기존 임상 증거와 연관 짓습니다. 머신러닝 모델에 대한 주요 비판 중 하나는 "블랙박스" 특성입니다 [52]. 머신러닝 모델은 지난 10년 동안 더욱 복잡해지면서 입력 특징 간의 복잡한 관계를 모델링함으로써 성능이 향상되었습니다. 그러나 복잡성이 증가함에 따라 해석 가능성과 설명력이 저하되었습니다. 의료와 같은 많은 사용 사례에서 의사 결정의 투명성 부족은 매우 중요하며, 모델 결과의 근거는 신뢰, 책임 및 의사 결정에 필수적입니다 [52]. XAI에서는 복잡한 머신러닝 모델의 설명력을 향상시키기 위해 다양한 기법이 제안되었습니다. SHAP는 예측을 '보상'으로, 기능의 가치를 '플레이어'로 가정하는 연합 게임 개념에서 파생된 XAI 프레임워크 유형입니다 [53].

Shapley 값을 계산하고 특징이 출력에 어떻게 영향을 미쳤는지 보여줌으로써 예측을 설명합니다. ML 알고리즘의 효과에 영향을 미친 각 특징의 중요도는 식 (2)에 나와 있는 대로 계산됩니다.

$$\phi(f, x) = \sum_{S \subseteq N_i} \frac{|S|! (N - |S| - 1)!}{|N|!} [f(x_S \cup \{x_i\}) - f(x_S)]$$

여기서 f 는 예측 모델을 나타내고, x 는 입력 특징 벡터이며, N 은 특징 집합, S 는 특징 집합의 부분 집합입니다.

특징 i 를 제외한 특징 벡터 x_S 는 XAI 분석 대상 특징이며, 여기서 x_S 는 부분집합 S 의 특징들을 기준값으로 대체한 특징 벡터이고, x_i 는 입력 벡터 x 에서 특징 i 의 값입니다. XAI 분석에서는 가장 우수한 성능을 보인 모델로 XGBOOST 모델 만을 사용했습니다.

LIME은 XAI 분야에서 널리 사용되는 프레임워크로, 머신러닝 예측 결과를 더 쉽게 이해할 수 있도록 도와줍니다. LIME은 특정 인스턴스의 로컬 컨텍스트에서 모델의 예측을 단순화하여 설명합니다. 이를 위해 해당 인스턴스를 중심으로 입력 데이터를 약간씩 변경하고, 이러한 변화가 모델의 예측에 어떤 영향을 미치는지 분석합니다. 이 과정을 통해 해당 인스턴스 주변에서 가장 영향력 있는 특징들을 파악할 수 있습니다. 각 특징의 중요도를 판단하기 위해 LIME은 가중치가 부여된 데이터 포인트들을 기반으로 간단하고 해석하기 쉬운 모델(예: 선형 회귀 모델)을 구축합니다. 이 가중치는 변경된 데이터 포인트가 원래 인스턴스에 얼마나 가까운지에 따라 할당됩니다.

이진 분류를 위한 XAI

두 개의 주요 SHAP 플롯 (그림 4 및 5)은 각 특징이 전체 데이터 세트에서 모델의 결정 경계에 미치는 영향을 보여줍니다. 그림 4는 특징별 SHAP 값의 평균 크기(평균 절대 SHAP)를 나타내며, 가장 영향력 있는 예측 변수를 강조합니다. 그림 5는 각 점이 특정 샘플의 SHAP 값에 해당하는 점을 나타낸 요약 그래프입니다. 오른쪽에 있는 점(양의 SHAP)은 분류를 "우울증" 쪽으로, 왼쪽에 있는 점(음의 SHAP)은 "비우울증" 쪽으로 기울게 합니다. 색상(빨간색에서 파란색)은 실제 특징 값을 나타냅니다(예: 높은 PSD 평균은 빨간색, 낮은 PSD 평균은 파란색).

그림 4. 이진 분류 특징 중요도.

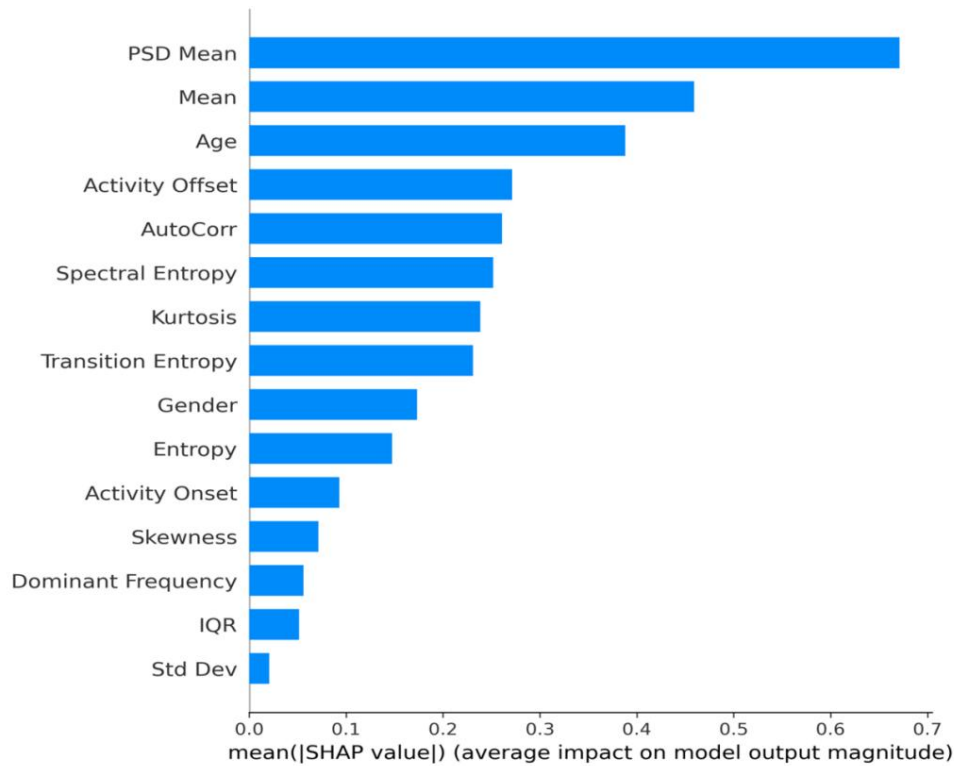
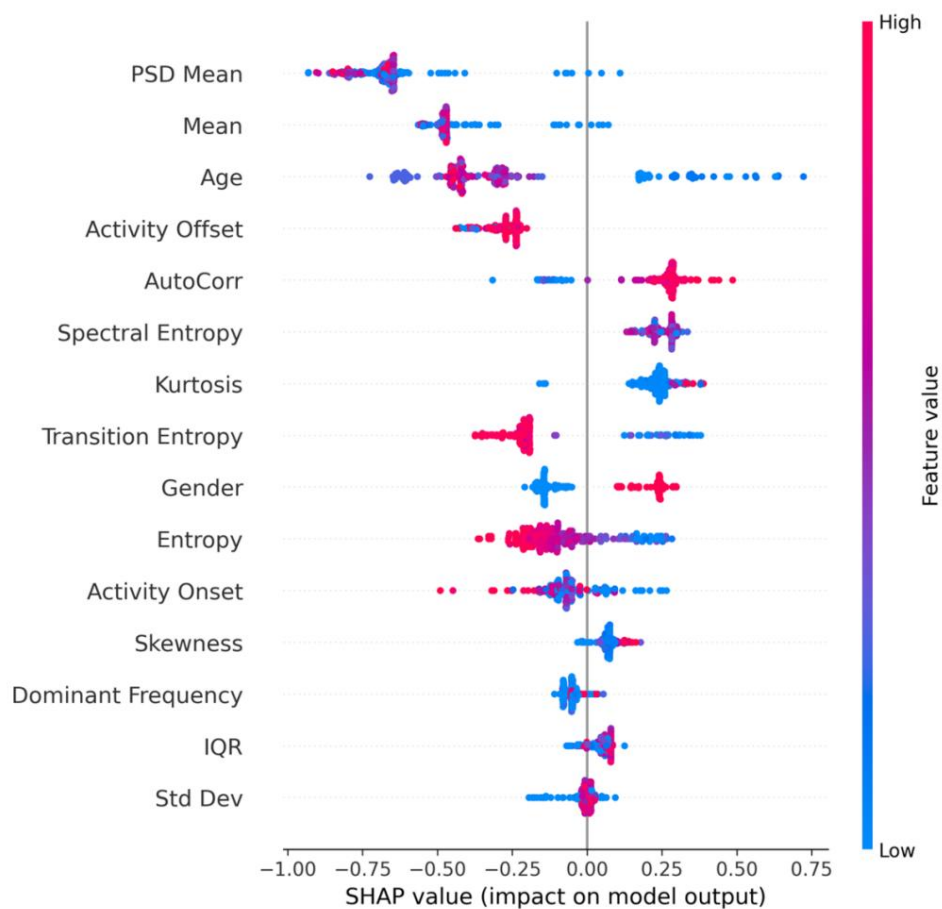


그림 5. 이진 분류에 대한 SHAP 요약 플롯.

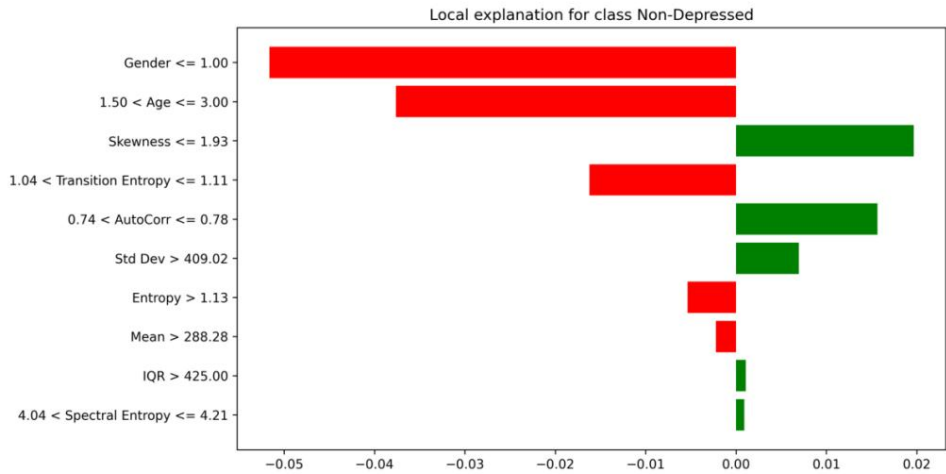


이 그래프들을 통해 다음과 같은 사실을 확인할 수 있습니다. PSD 평균값이 일관되게 가장 큰 평균적인 영향을 미치는 것으로 나타났습니다. PSD 평균값이 클수록(활동 데이터에서 저주파수 파워가 강함을 반영하는 것으로 추정됨) 우울증 발생 가능성이 높은 것으로 나타났습니다. 평균 활동 수준과 연령도 비슷한 영향을 미치는 것으로 보아 인구통계학적 요인(예: 고령자)과 전반적인 일일 활동 수준이 분류에 상당한 영향을 미치는 것으로 나타났습니다. 활동 시간차와 자기상관도 두드러지게 나타났는데, 이는 불규칙하거나 경직된 일주기 리듬(낮은 시간차와 높은 자기상관)이 우울증 행동과 관련이 있다는 임상 연구 결과와 일치합니다.

PSD 평균은 일반적으로 가장 큰 평균적인 영향을 미칩니다. PSD 평균값이 클수록(활동 데이터에서 저주파수 전력이 더 강함을 반영하는 것일 수 있음) 우울증 발생 가능성이 더 높은 것으로 나타났습니다.

평균 활동 수준과 연령은 밀접한 관계를 보이는데, 이는 인구 통계학적 요인(예: 고령자)과 전반적인 일일 활동 수준이 분류에 상당한 영향을 미친다는 것을 나타냅니다.

그림 6. 우울증이 없는 집단에 대한 지역적 설명.



SHAP와 LIME은 이론적으로 서로 다르지만(SHAP는 게임 이론적 Shapley 값을 사용하고, LIME은 국소 선형 근사치를 사용함), 둘 다 유사한 영향력 있는 특징들로 수렴하여 모델이 학습한 표현에 대한 신뢰도를 높입니다. 특히, PSD 평균, 평균 활동 수준, 그리고 일주기 리듬 지표(활동 오프셋, 자동 보정 등)가 가장 강력한 예측 변수로 반복적으로 나타납니다.

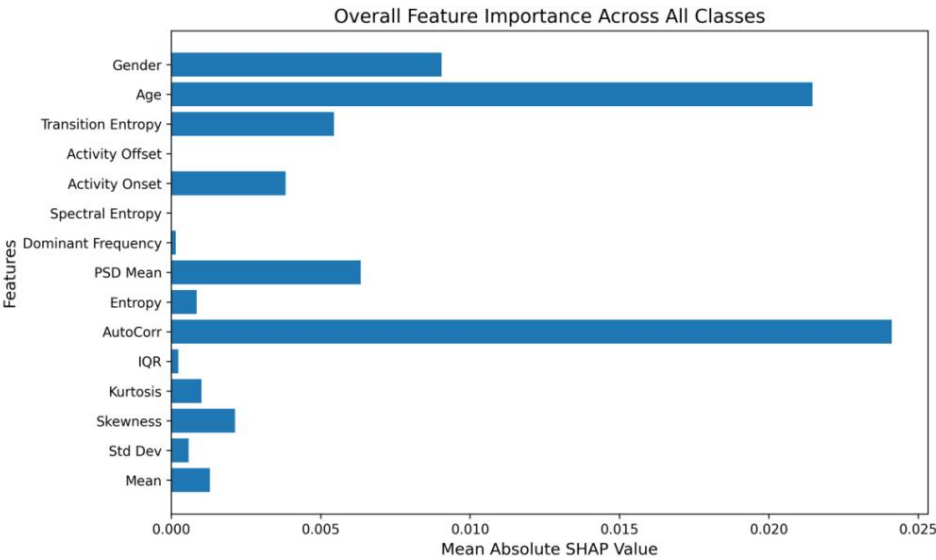
다중 클래스 분류를 위한 XAI

우리는 설명 가능한 AI 접근 방식을 "정상 사례", "경미한 우울증"으로 구분하는 3개 클래스 설정으로 확장했습니다.

그림 6의 단일 테스트 사례를 설명하기 위해 LIME을 적용했습니다. LIME은 이 사례 주변의 작은 교란이 우울증 예측 확률에 어떤 영향을 미치는지 보여주는 국소적으로 선형적인 "규칙"을 생성합니다. 빨간색 막대는 사례를 "우울증" 쪽으로, 녹색 막대는 "비우울증" 쪽으로 밀어냅니다. 예를 들어, "성별≤1.0"(여성) 또는 "나이"가 많을수록 다른 활동 관련 특징들도 일치할 경우 모델의 우울증 진단 가능성이 높아질 수 있습니다. LIME은 각 조건에 대한 수치 가중치를 제공하여 자기상관, 엔트로피 또는 왜곡도의 임계값이 국소적인 결정 경계를 어떻게 형성하는지 보여줍니다. 이러한 규칙은 SHAP 결과와 잘 일치합니다(예: 높은 자기상관은 우울증 위험을 나타내는 경우가 많음).

그리고 “중등도 우울증”을 진단 기준으로 삼았습니다. SHAP는 전반적인 특징 중요도(클래스별 및 전체 특징 중요도)를 파악하는 데 사용되었고, LIME은 개별 사례별 특징 중요도를 설명하는 데 사용되었습니다. 그림 7은 세 가지 클래스 모두에서 집계된 평균 절대 SHAP 값을 보여주며, 모델 예측에 일관되게 영향을 미치는 주요 특징들을 강조합니다. 특히, 자동 수정(AutoCorr)과 연령이 주요 요인으로 나타나는데, 이는 반복적이거나 경직된 활동 패턴을 보이는 고령층이 우울증 위험이 더 높다는 임상적 관찰 결과와 일치합니다.

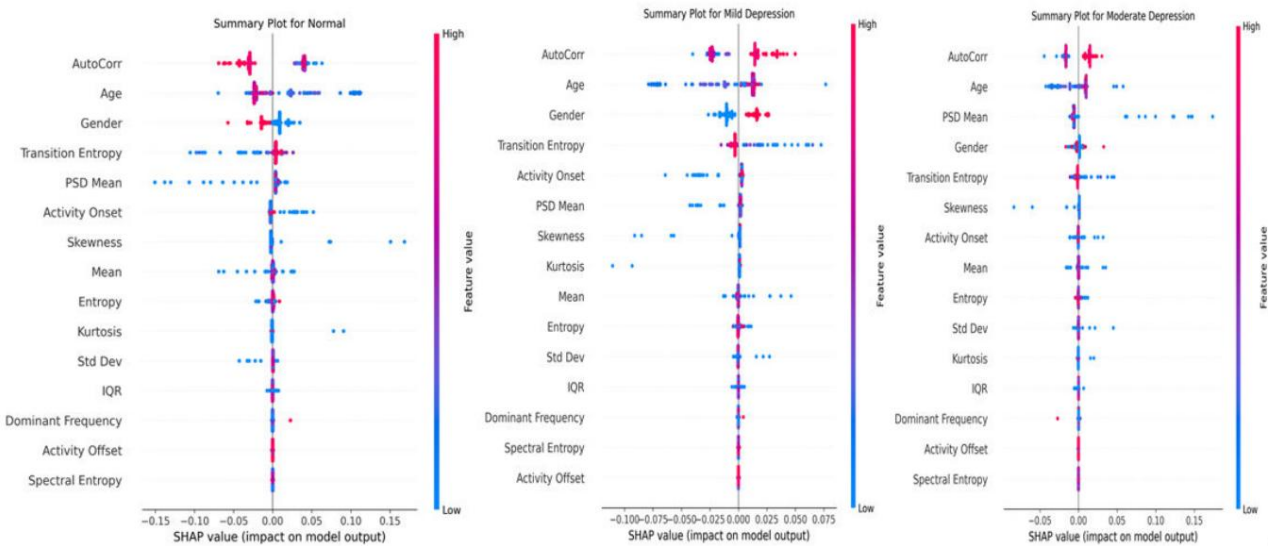
그림 7. 다중 클래스 특징 중요도.



각 클래스에 대한 보다 심층적인 분석을 위해 별도의 SHAP 요약 플롯을 생성했습니다. 정상 (그림 8A): 낮은 AutoCorr 값은 일반적으로 정상 클래스를 나타내며, 이는 보다 유연하거나 다양한 일상 리듬(반복적인 활동이 적음)을 시사합니다.

우울증이 없는 상태와 일치합니다. 연령과 PSD 평균값이 중간 정도일 경우 예측 결과가 경증 또는 중등도 우울증에서 멀어집니다.

그림 8. 다중 클래스 분류에 대한 SHAP 요약 플롯.



경미한 우울증 (그림 8B): 경미한 사례에서는 AutoCorr와 전이 엔트로피의 기여도가 높게 나타납니다. 많은 경우, 적당한 일주기 리듬과 활동의 주파수 구성의 미묘한 변화(PSD 평균으로 포착됨)가 결합되어 모델이 경미한 우울증 쪽으로 기울어집니다.

경미한 우울증을 나타내는 반면, 높은 자기상관성과 고령은 종종 중등도 우울증을 나타냅니다. 이러한 구분은 동일한 특징이라도 실제 범위가 다른 특징과의 상호작용에서 나타나는 미묘한 차이에 따라 계층별로 다른 영향을 미칠 수 있음을 강조합니다.

중등도 우울증 (그림 8C): 이 유형은 종종 더 심각하거나 고착화된 불규칙성을 나타냅니다. 연령, 자동 수정(AutoCorr), 그리고 높은 PSD 평균값은 해당 사례를 경증에서 중등도 수준으로 악화시키는 요인으로 작용하는 것으로 보입니다. 여기서는 인구통계학적 요인과 일주기 리듬의 교란이 더욱 강하게 얽혀 있으며, 이는 더 심각한 우울증 상태를 반영합니다.

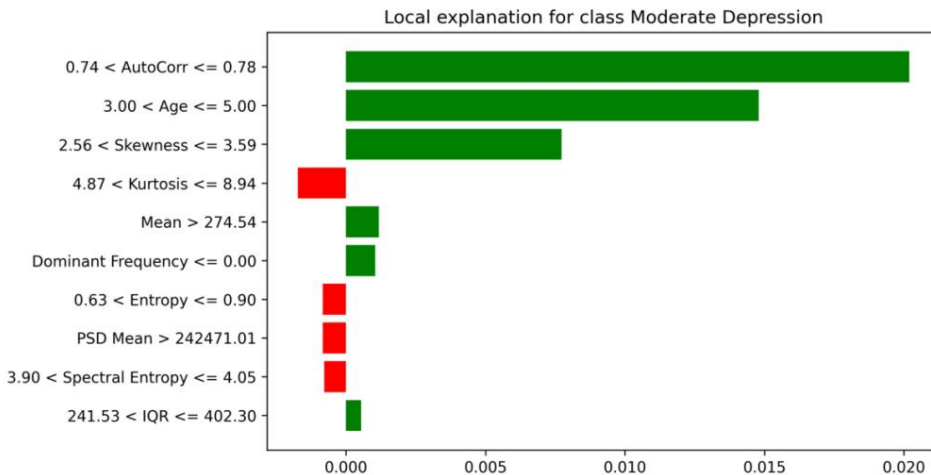
SHAP의 클래스별 결과를 인스턴스 수준에서 검증하기 위해 단일 테스트 예제에 LIME을 사용했습니다. 그림 9는 "중등도 우울증" 클래스에 대한 모델의 설명을 보여줍니다. LIME은 임계값 기반 "규칙"(예: $0.74 < \text{AutoCorr} \leq 0.78$)을 생성하여 클래스 멤버십 확률을 국소적으로 조정합니다. 녹색 막대는 모델이 "중등도 우울증"을 선택하는 방향으로 기울어짐을 나타냅니다. 예를 들어, AutoCorr 범위가 0.74를 초과하거나 나이가 특정 임계값을 초과하면 이 클래스의 확률이 상당히 높아질 수 있습니다. 빨간색 막대는

모든 클래스에서 AutoCorr는 일관되게 가장 중요한 기능 중 하나로 꼽힙니다. 하지만 구체적인 영향력은 클래스에 따라 다릅니다. 적당한 자기상관은 사례를 정상적인 상태에서 다른 상태로 이끌 수 있습니다.

막대는 예측값을 "중등도 우울증"에서 멀어지게 하며, 이는 예측값과 충돌하는 특징이나 임계값을 나타냅니다.

중등도 심각도의 전형적인 프로필(예: 낮은 PSD 평균 또는 높은 침도).

그림 9. 중등도 우울증 유형에 대한 LIME 설명.



이러한 국소적 관점은 SHAP 관찰 결과와 일치합니다. 즉, 일주기 관련 변수(예: 자동 수정 및 전환 엔트로피)와 인구 통계학적 요인(예: 연령)이 핵심적인 역할을 합니다. 경미한 경우에는 이러한 지표들이 부분적으로만 교란된 것처럼 보일 수 있지만, 중등도의 경우에는 이러한 지표들이 더 심각한 우울증 증상과 관련된 범위에서 지속적으로 나타납니다.

XGBoost에 대한 설명만 제공했다면 놓쳤을 통찰력입니다.

다중 클래스 분류의 경우, 평균 SHAP 값을 살펴보면 XGBoost와 신경망이 다중 클래스 분류에서 부분적으로 겹치지만 궁극적으로는 서로 다른 정보에 의존하며, 그 근거를 부여하는 척도 또한 현저히 다르다는 것을 알 수 있습니다. XGBoost의 가장 강력한 신호는 시간 구조 및 인구 통계학적 변수에서 비롯되는데, 자동 보정(AutoCorr)이 가장 큰 영향을 미치고, 그 다음으로 연령이 뒤를 잇으며, 성별과 전환 엔트로피 또한 상당한 영향을 줍니다. 반면 활동 오프셋과 같은 특징은 거의 영향을 미치지 않습니다.

다양한 설명 비교하기 XGBoost와 신경망 모델

XGBoost가 가장 높은 정확도를 보였지만, 설명력이 모델에 따라 달라지는지 확인하기 위해 인공신경망(ANN)에 대한 XAI 분석을 반복 수행했습니다.

이진 분류의 경우, 두 모델 모두 PSD 평균을 우울증 여부 판단의 주요 요인으로 꼽아 해당 특징의 견고성을 확인했습니다. 그 외의 설명은 서로 달랐습니다. XGBoost는 평균, 나이, 활동 종료 시점, 자동 보정 등 약 10개의 추가 변수에 중간 정도의 중요도를 분산시킨 반면, ANN은 나머지 거의 모든 기여도를 활동 시작 시점, 침도, 활동 종료 시점의 세 가지 특징에 집중시켰습니다.

이 비교는 두 가지 방식으로 해석 가능성을 강화합니다. PSD 평균에 대한 모델 간 일치는 이를 일관된 생리적 지표로 검증하며, 서로 다른 이차 패턴은 구조별 민감도를 드러내어 XGBoost만을 사용한 설명으로는 얻을 수 없는 보완적인 통찰력을 제공합니다.

이진 분류의 경우, 두 모델 모두 PSD 평균을 우울증 유무를 판단하는 주요 요인으로 꼽아 해당 특징의 견고성을 확인했습니다. 그 외의 설명은 서로 달랐습니다. XGBoost는 평균, 나이, 활동 종료 시점, 자동 보정 등 약 10개의 추가 변수에 중간 정도의 중요도를 부여한 반면, ANN은 나머지 거의 모든 기여도를 단 3개의 특징(활동 시작 시점, 침도, 활동 종료 시점)에 집중시켰습니다. 이러한 비교는 두 가지 측면에서 해석 가능성을 높입니다. (1) PSD 평균에 대한 모델 간 일치는 이를 일관된 생리적 지표로 검증하고, (2) 서로 다른 이차 패턴은 아키텍처별 민감도를 드러내어 상호 보완적인 설명을 제공합니다.

인공신경망(ANN)은 거의 정반대의 결과를 보여줍니다. ANN에서 가장 영향력 있는 두 가지 단서는 고차 통계량인 침도와 활동 오프셋인 반면, XGBoost에서 지배적인 변수들은 노이즈 폴러어로 축소됩니다. 그럼에도 불구하고 두 모델 모두 PSD 평균을 관련 지표로 수렴하는데, 이는 XGBoost에서는 가중치가 미미하고 ANN에서는 상대적으로 두드러지지만, 해당 스펙트럼 특징의 생리학적 타당성에 대한 확신을 높여줍니다. 또한 ANN의 절대적인 SHAP 크기는 트리 양상 불보다 대략 두 자릿수 정도 작다는 점에 주목할 필요가 있습니다. 이는 네트워크가 학습한 더 부드러운 결정 표면과 KernelExplainer가 생성하는 보수적인 추정치를 반영합니다. 따라서 비교는 다른 스케일 간의 비교가 아닌 각 모델의 스케일 내에서 이루어져야 합니다. 종합적으로 볼 때, PSD 평균에 대한 일치와 다른 부분에서의 차이는 두 아키텍처가 상호 보완적인 패턴을 추출한다는 것을 보여줍니다. 트리 구조는 이질적인 시간적-인구통계학적 상호작용을 활용하는 반면, 네트워크 구조는 데이터를 분포 형태 지표로 정제하여 우울증 심각도가 어떻게 인코딩되는지에 대한 더욱 풍부하고 교차 검증된 그림을 제공합니다.

XAI 분석 결과에 대한 논의

SHAP-LIME 분석을 종합적으로 수행한 결과 몇 가지 일관된 주제가 드러났습니다.

일주기 리듬 및 활동 측정

본 모델이 자기상관 및 스펙트럼 활동 특징에 중점을 둔 것은 일주기 리듬 장애와 기분 장애를 연관시키는 많은 연구 결과와 일치합니다. 예를 들어, 임상 및 실험적 증거에 따르면 일주기 리듬의 둔화 또는 변화는 우울증과 밀접한 관련이 있습니다 [54,55]. 웨어러블 기기를 사용한 최근의 중단 연구에서는 일주기 위상 변화가 우울증 환자의 기분 증상 악화에 선행하는 경우가 많다는 사실이 밝혀졌습니다 [56]. 또한, 일주기 리듬을 재조정하는 치료법(밝은 빛 치료, 사회적 리듬 치료 등)은 우울증 증상을 효과적으로 완화시켜 [55], 리듬 장애가 원인일 가능성을 시사합니다. 그러나 기분과 일주기 과정은 양방향으로 연결되어 있으므로 [55,57], 변화된 활동 패턴은 우울증 상태에 영향을 미치기도 하고 우울증 상태를 반영하기도 합니다. 요컨대, 본 연구 결과는 기존에 알려진 연관성과 일치하지만, 이러한 리듬 관련 특징이 우울증을 유발한다는 것을 그 자체로 증명하는 것은 아닙니다.

연령 및 인구통계학적 영향

우울증 위험의 연령 및 성별 차이는 잘 알려져 있으며 [57,58], 우리의 모델은 이러한 패턴을 포착합니다. 예를 들어, 노년층은 수면-각성 주기가 더 빠르고 변동성이 적은 경향이 있는데 [59], 이는 노년기 우울증 위험과 상호작용할 수 있습니다. 역학 연구에 따르면 여성은 남성에 비해 주요 우울증 발병 확률이 약 두 배 높습니다 [58]. 본 모델의 성별 특성 추세도 이와 일치합니다. 그러나 다른 위험 요인과 마찬가지로 연령 및 성별의 영향은 다면적이며 종종 간접적입니다. 생물학적 노화는 건강, 생활 방식, 뇌 기능 등 여러 변화를 가져오고, 성별 차이는 호르몬, 사회심리적, 진단적 요인을 반영합니다. 따라서 연령 및 인구통계학적 특성이 우울증과 상관관계가 있지만, 그 기전적 경로는 여전히 추측에 불과합니다. 실제로 이러한 변수들은 복잡하고 상황에 따라 달라지는 방식으로 위험을 조절할 가능성이 높으므로 [57], 추가 연구 없이 인과관계를 단정짓는 것은 신중해야 합니다.

경증과 중증도의 구분

우울증

다중 클래스 SHAP 플롯은 경증에서 중증도 수준으로의 특정 전환이 일상 활동 불규칙성의 증가와 관련되는 경우가 많다는 것을 보여줍니다. 전환 엔트로피와

AutoCorr는 특히 의미심장한데, 이는 수면-각성 전환이 더 혼란스러울수록 우울증이 심해진다는 심리학 연구 결과와 유사합니다 [60]. 그럼에도 불구하고 이는 여전히 관찰에 의한 상관관계일 뿐입니다. 예를 들어, 불규칙한 수면이 우울증을 유발하는지 아니면 주로 생리적 영향만을 반영하는지는 불분명합니다. 이를 명확히 밝히기 위해서는 SHAP 및 LIME만으로는 보여줄 수 없는, 보다 구체적인 연구(예: 수면 스케줄 조작)가 필요합니다.

로컬 설명 값

인스턴스별 힘 플롯 또는 LIME 규칙을 사용하면 위험 요인의 임상적으로 관련된 분석을 통해 인구 통계학적 또는 일주기적 지표의 미묘한 변화가 환자의 분류를 "정상"에서 "경증" 또는 "중증도 우울증"으로 어떻게 바꿀 수 있는지 강조할 수 있습니다. 이러한 국소적 해석 가능성은 개인 맞춤형 의학과 상황 인식 진단의 필요성에 대한 문헌의 강조와 일치합니다 [61].

일반적으로 XAI 접근 방식은 데이터 기반 우울증 탐자가 일주기 리듬 및 인구 통계학적 특성에 크게 의존한다는 것을 확인시켜 주며, 이는 일상적인 행동 패턴과 정신 건강 상태 간의 연관성에 대한 기존 연구 결과와 일맥상통합니다. SHAP과 LIME은 전체적인(전역적) 관점과 개별 사례에 대한(지역적) 관점을 모두 제공함으로써 정신 건강 분야에서 머신러닝의 투명성과 신뢰성을 강화합니다. 이러한 연관성은 고무적이지만, 진정한 인과 관계를 확인하려면 SHAP과 LIME에서 밝혀진 상관관계를 넘어선 표적화된 중단 연구 및 중재 연구가 필요합니다.

본 연구에서는 웨어러블 활동량 측정 데이터를 기반으로 우울증을 탐지하고 중증도를 분류하기 위해 설명 가능한 인공지능(AI)을 접목한 머신러닝 프레임워크를 제시했습니다. XGBoost를 ADASYN 및 특징 선택 알고리즘과 결합하여 다른 방법들에 비해 일관되게 우수한 성능을 달성함을 입증했습니다. SHAP 및 LIME 설명 분석 결과, 일주기 리듬 관련 요인과 인구통계학적 요인(예: PSD 평균, 연령, 자동 보정)이 우울증 상태 식별에 중요한 역할을 하는 것으로 나타났습니다. 제안된 시스템은 전역적 및 지역적 해석 가능성을 모두 제공함으로써 임상가와 연구자들이 모델의 판단을 이해하고 신뢰할 수 있도록 지원하여, 보다 효과적인 치료 계획을 수립하고 정신 건강 관리 분야에서 AI 기반 도구의 활용도를 높이는 데 기여할 수 있습니다.

데이터 가용성 데이터 세트는 Depresjon Dataset [62]에서 다운로드할 수 있습니다.

아래 충돌 신고된 사항 없음.

멀티미디어 부록 1 모델의 하이퍼파라미터.
[DOCX 파일(Microsoft Word 파일), 14 KB-멀티미디어 부록 1]

참고 자료

1. Santomauro DF, Mantilla Herrera AM, Shadid J, et al. 2020년 COVID-19 팬데믹으로 인한 204개 국가 및 지역의 우울증 및 불안 장애의 전 세계적 유병률 및 부담. The Lancet. 2021년 11월;398(10312):1700-1712. [doi: 10.1016/S0140-6736(21)02143-7]

2. Gulland A. NHS 조사에 따르면 여성의 정신 질환 발병률이 남성보다 높습니다. BMJ. 2016;i5320. [doi: [10.1136/bmj.i5320](https://doi.org/10.1136/bmj.i5320)]
3. Varma P, Junge M, Meaklim H, Jackson ML. 젊은 사람들은 COVID-19 팬데믹 기간 동안 스트레스, 불안 및 우울증에 더 취약합니다: 글로벌 횡단면 조사. Prog Neuropsychopharmacol Biol Psychiatry. 2021년 7월 13일;109:110236. [doi: [10.1016/j.pnpbp.2020.110236](https://doi.org/10.1016/j.pnpbp.2020.110236)] [메드라인: [33373680](https://pubmed.ncbi.nlm.nih.gov/33373680/)]
4. 세계보건기구. 우울 장애(우울증) - who.int. 2023. URL: <https://www.who.int/news-room/자료집/상세정보/우울증> [접근일: 2025년 9월 4일]
5. 세계보건기구. 자살 - 인도 - who.int. URL: <https://www.who.int/india/health-topics/suicide> [접근됨] 2025년 9월 4일]
6. Chand SP, Arif H. 우울증. StatPearls Publishing; 2024. URL: <https://www.ncbi.nlm.nih.gov/books/NBK430847/> [접근일: 2025년 9월 4일]
7. Tolentino JC, Schmidt SL. DSM-5 진단 기준과 우울증 심각도: 임상 진료에 대한 함의. Front Psychiatry. 2018;9:450. [doi: [10.3389/fpsy.2018.00450](https://doi.org/10.3389/fpsy.2018.00450)] [메드라인: [30333763](https://pubmed.ncbi.nlm.nih.gov/30333763/)]
8. McKee J, Brahm N. 의학적 모방: 정신과적 증상에 대한 감별 진단 고려 사항. 정신 건강 임상. 2016년 11월;6(6):289-296. [doi: [10.9740/mhc.2016.11.289](https://doi.org/10.9740/mhc.2016.11.289)] [메드라인: [29955484](https://pubmed.ncbi.nlm.nih.gov/29955484/)]
9. Currier MB, Nemeroff CB. 우울증은 암의 위험 요인이다: 병태생리학적 발전에서 치료적 함의까지. Annu Rev Med. 2014;65(1):203-221. [doi: [10.1146/annurev-med-061212-171507](https://doi.org/10.1146/annurev-med-061212-171507)] [메드라인: [24215332](https://pubmed.ncbi.nlm.nih.gov/24215332/)]
10. Liu H, Luiten PGM, Eisel ULM, De Jongste MJL, Shoemaker RG. 심근경색 후 우울증: TNF- α -혈뇌장벽의 유도된 변화와 그에 따른 치료적 함의. 신경과학 및 생물행동 리뷰. 2013년 5월;37(4):561-572. [doi: [10.1016/j.neubiorev.2013.02.004](https://doi.org/10.1016/j.neubiorev.2013.02.004)]
11. Kohn R, Saxena S, Levav I, Saraceno B. 정신 건강 관리의 치료 격차. 세계보건기구 회보. 11월 2004;82(11):858-866. [Medline: [15640922](https://pubmed.ncbi.nlm.nih.gov/15640922/)]
12. Sharp LK, Lipsky MS. 생애주기 전반에 걸친 우울증 선별검사: 일차 진료 환경에서 사용하기 위한 측정 도구 검토. Am Fam Physician. 2002년 9월 15일;66(6):1001-1008. [Medline: [12358212](https://pubmed.ncbi.nlm.nih.gov/12358212/)]
13. Montgomery SA, Asberg M. 변화에 민감하도록 설계된 새로운 우울증 척도. Br J Psychiatry. 1979년 4월;134(4):382-389. [doi: [10.1192/bjp.134.4.382](https://doi.org/10.1192/bjp.134.4.382)] [메드라인: [444788](https://pubmed.ncbi.nlm.nih.gov/444788/)]
14. Quilty LC, Robinson JJ, Rolland JP, Fruyt FD, Rouillon F, Bagby RM. 몽고메리-아스베르크 해저 터널의 구조 우울증 치료 과정 중 우울증 평가 척도. Int J Methods Psych Res. 2013년 9월;22(3):175-184. [doi: [10.1002/mpr.1388](https://doi.org/10.1002/mpr.1388)]
15. Herrmann N, Black SE, Lawrence J, Szekely C, Szalai JP. 씨니브룩 뇌졸중 연구: 전향적 연구 우울 증상 및 기능적 결과. 뇌졸중. 1998년 3월;29(3):618-624. [doi: [10.1161/01.str.29.3.618](https://doi.org/10.1161/01.str.29.3.618)] [메드라인: [9506602](https://pubmed.ncbi.nlm.nih.gov/9506602/)]
16. Remes O, Mendes JF, Templeton P. 우울증의 생물학적, 심리적, 사회적 결정 요인: 최근 문헌 검토. Brain Sci. 2021년 12월 10일;11(12):1633. [doi: [10.3390/brainsci11121633](https://doi.org/10.3390/brainsci11121633)] [메드라인: [34942936](https://pubmed.ncbi.nlm.nih.gov/34942936/)]
17. Smith KM, Renshaw PF, Bilello J. 우울증 진단: 현재 및 새로운 방법. Compr Psychiatry. 2013년 1월;54(1):1-6. [doi: [10.1016/j.comppsy.2012.06.006](https://doi.org/10.1016/j.comppsy.2012.06.006)]
18. Richter T, Fishbain B, Markus A, Richter-Levin G, Okon-Singer H. 기계 학습 기반 분석을 사용하여 불안과 우울증의 행동적 차이. Sci Rep. 2020;10(1):16381. [doi: [10.1038/s41598-020-72289-9](https://doi.org/10.1038/s41598-020-72289-9)]
19. Muhammad D, Ahmed I, Ahmad MO, Bendeache M. 효율적인 의료 진단을 위한 무작위화된 설명 가능한 기계 학습 모델. IEEE J Biomed Health Inform. 2024년 11월 13일;PP:1-10. [doi: [10.1109/JBHI.2024.3491593](https://doi.org/10.1109/JBHI.2024.3491593)] [메드라인: [40030196](https://pubmed.ncbi.nlm.nih.gov/40030196/)]
20. Topol EJ. 고성능 의학: 인간과 인공지능의 융합. Nat Med. Jan 2019;25(1):44-56. [doi: [10.1038/s41591-018-0300-7](https://doi.org/10.1038/s41591-018-0300-7)] [메드라인: [30617339](https://pubmed.ncbi.nlm.nih.gov/30617339/)]
21. Abd-Alrazaq A, AlSaad R, Aziz S, et al. 불안 및 우울증을 위한 웨어러블 인공지능: 범위 검토. J Med Internet Res. 2023년 1월 19일;25:e42672. [doi: [10.2196/42672](https://doi.org/10.2196/42672)] [메드라인: [36656625](https://pubmed.ncbi.nlm.nih.gov/36656625/)]
22. Paul N, Juliet S. 정신 건강 예측을 위한 기계 학습 기술의 비교 분석. 발표: 2023년 제8회 국제 통신 및 전자 시스템 컨퍼런스(ICCES); 인도 코임바토르. 2023.[doi: [10.1109/ICCES57224.2023.10192763](https://doi.org/10.1109/ICCES57224.2023.10192763)]
23. Zhang J, Richardson JD, Dunkley BT. 자기파와 연결체와 기계 학습을 이용한 외상 후 스트레스 장애 분류. Sci Rep. 2020;10(1):5937. [doi: [10.1038/s41598-020-62713-5](https://doi.org/10.1038/s41598-020-62713-5)]
24. Mikolas P, Vahid A, Bernardoni F, et al. 의료 기록의 실제 임상 데이터를 기반으로 ADHD를 식별하는 기계 학습 분류기 훈련. Sci Rep. 2022년 7월 28일;12(1):12934. [doi: [10.1038/s41598-022-17126-x](https://doi.org/10.1038/s41598-022-17126-x)] [메드라인: [35902654](https://pubmed.ncbi.nlm.nih.gov/35902654/)]

25. Spulber S, Elberling F, Svensson J, Tiger M, Ceccatelli S, Lundberg J. 주요 우울 장애 환자의 활동 패턴은 증상 심각도와 상관관계가 있다. *Transl Psychiatry*. 2022년 6월 2일;12(1):226. [doi: [10.1038/s41398-022-01989-9](https://doi.org/10.1038/s41398-022-01989-9)] [메드라인: [35654778](https://pubmed.ncbi.nlm.nih.gov/35654778/)]
26. Garcia-Ceja E, Riegler M, Jakobsen P, et al. 단극성 및 양극성 우울증 환자의 운동 활동 기반 분류. 2018 IEEE 제31회 국제 컴퓨터 기반 의료 시스템 심포지엄 (CBMS) 발표. 칼스타드. 2018.[doi: [10.1109/CBMS.2018.00062](https://doi.org/10.1109/CBMS.2018.00062)]
27. Frogner JI, Noori FM, Halvorsen P, 외. 운동 활동에 대한 1차원 합성곱 신경망
우울증 진단에서의 측정. 발표: MM '19; 9-15; 프랑스 니스. 2019년 10월 15일. URL : <https://dl.acm.org/doi/proceedings/10.1145/3347444> [doi: [10.1145/3347444.3356238](https://doi.org/10.1145/3347444.3356238)]
28. Jakobsen P, Garcia-Ceja E, Riegler M, et al. 우울증 환자의 운동 활동 시계열에 기계 학습 적용
건강한 대조군과 비교한 양극성 장애 및 단극성 장애 환자. *PLoS ONE*. 2020;15(8):e0231995. [doi: [10.1371/journal.pone.0231995](https://doi.org/10.1371/journal.pone.0231995)] [메드라인: [32833958](https://pubmed.ncbi.nlm.nih.gov/32833958/)]
29. Espino-Salinas CH, Galván-Tejada CE, Luna-García H 등. 2차원 컨벌루션 신경망
우울증 데이터 세트의 운동 활동 시계열을 이용한 실시간 우울증 에피소드 감지. *생명공학(바젤)*.
2022년 9월 9일;9(9):458. [doi: [10.3390/bioengineering9090458](https://doi.org/10.3390/bioengineering9090458)] [메드라인: [36135004](https://pubmed.ncbi.nlm.nih.gov/36135004/)]
30. Li Y. 기계 학습을 활용한 정신 건강 장애 예측 및 특징 중요도 해석. 발표:
2023 제3회 국제 컴퓨터 기술 및 정보 과학 심포지엄(ISCTIS). 257-261; 2023.[doi: [10.1109/ISCTIS58954.2023.10213032](https://doi.org/10.1109/ISCTIS58954.2023.10213032)]
31. Malik A, Shabaz M, Asenso E. 코로나19 위기 동안 우울증을 감지하기 위한 기계 학습 기반 모델. *Sci Afr*.
2023년 7월;20:e01716. [doi: [10.1016/j.sciaf.2023.e01716](https://doi.org/10.1016/j.sciaf.2023.e01716)] [메드라인: [37214195](https://pubmed.ncbi.nlm.nih.gov/37214195/)]
32. Handing EP, Strobl C, Jiao Y, Feliciano L, Aichele S. 유럽의 중년 및 노년 남녀의 우울증 예측인자: 기계 학습 접근법. *Lancet Reg Health Eur*. 2022년 7
월;18:100391. [doi: [10.1016/j.lanepe.2022.100391](https://doi.org/10.1016/j.lanepe.2022.100391)] [메드라인: [35519235](https://pubmed.ncbi.nlm.nih.gov/35519235/)]
33. Haque UM, Kabir E, Khanam R. 기계 학습 방법을 이용한 아동 우울증 탐지. *PLoS ONE*.
2021;16(12):e0261131. [doi: [10.1371/journal.pone.0261131](https://doi.org/10.1371/journal.pone.0261131)] [메드라인: [34914728](https://pubmed.ncbi.nlm.nih.gov/34914728/)]
34. Pacheco-González SL, Zanella-Calzada LA, Galván-Tejada CE, Chávez-Lamas NM, Rivera-Gómez JF, Galván-Tejada JI. 우울증 에피소드 탐지를 위한 5
가지 분류기 평가. *RCS*. 2019;148(10):129-138. [doi: [10.13053/rcs-148-10-11](https://doi.org/10.13053/rcs-148-10-11)]
35. Galván-Tejada CE, Zanella-Calzada LA, Gamboa-Rosales H, et al. 단극성 및 양극성 환자의 우울증 에피소드 탐지: 활동 동작 신호를 정보원으로 사용하는 유전 알
고리즘을 이용한 특징 추출 및 특징 선택 방법론. 모바일 정보 시스템. 2019년 4월 23일;2019(1):1-12. [doi: [10.1155/2019/8269695](https://doi.org/10.1155/2019/8269695)]
36. Garcia-Ceja E, Riegler M, Jakobsen P, et al. Depresjon: 단극성 및 양극성 장애 환자의 우울증 삽화에 대한 운동 활동 데이터베이스. 2018년 멀티미디어 시스템 컨퍼
런스 발표; 암스테르담. 2018.[doi: [10.1145/3204949.3208125](https://doi.org/10.1145/3204949.3208125)]
37. Garcia EA, He H, Bai Y, Li SA ADASYN: 불균형 학습을 위한 적응형 합성 샘플링 접근법. 발표됨
2008 IEEE 국제 신경망 공동 학술회의(IEEE 세계 컴퓨팅 지능 학술회의); 1322-1328; 홍콩. 2008.[doi: [10.1109/IJCNN.2008.4633969](https://doi.org/10.1109/IJCNN.2008.4633969)]
38. Zanella-Calzada LA, Galván-Tejada CE, Chávez-Lamas NM, et al. 운동 활동 신호의 특징 추출: 단극성 및 양극성 환자의 우울증 에피소드 감지를 향하여.
Diagnostics (Basel). 2019년 1월 10일;9(1):8. [doi: [10.3390/diagnostics9010008](https://doi.org/10.3390/diagnostics9010008)] [메드라인: [30634621](https://pubmed.ncbi.nlm.nih.gov/30634621/)]
39. Maczák B, Gingl Z, Vadai G. 인간 활동의 일반적인 스펙트럼 특성 및 그에 내재된 스케일 프리 변동.
Sci Rep. 2024년 1월 31일;14(1):2604. [doi: [10.1038/s41598-024-52905-8](https://doi.org/10.1038/s41598-024-52905-8)] [메드라인: [38297022](https://pubmed.ncbi.nlm.nih.gov/38297022/)]
40. Wilson CE, Crouse JJ, Carpenter JS, Hickie IB. 기분 장애의 일주기적 요소: 수면-각성 주기,
생체 리듬 및 시간치료. Kushida CA 편집. 수면 및 일주기 리듬 백과사전. 2판. Academic Press; 2023:766-779. [doi: [10.1016/B978-0-12-822963-7.00147-X](https://doi.org/10.1016/B978-0-12-822963-7.00147-X)]
ISBN: 978-0-323-91094-1 41. Langensiepen C, Lotfi A. 주변 지능 환경에서의 불확실성 측정. 2017 IEEE 국제 퍼지 시스템 학회(FUZZ-IEEE) 발
표. 1-5; 2017.[doi: [10.1109/FUZZ-IEEE.2017.8015765](https://doi.org/10.1109/FUZZ-IEEE.2017.8015765)]
42. Hosmer Jr DW, Lemeshow S, Sturdivant RX. 응용 로지스틱 회귀 분석. John Wiley & Sons; 2013. [doi: [10.1002/9781118548387](https://doi.org/10.1002/9781118548387)]
43. Do TN, Poulet F. 대규모 이미지 데이터셋 분류를 위한 병렬 다중 클래스 로지스틱 회귀. In: *Advanced Computational Methods for Knowledge Engineering Advances in Intelligent Systems and Computing*. Springer; 2015:255-266. [doi: [10.1007/978-3-319-17996-4_23](https://doi.org/10.1007/978-3-319-17996-4_23)]
44. Chandra MA, Bedi SS. SVM 및 이미지 분류에서의 응용에 대한 조사. *Int j inf tecnol*. 10월
2021;13(5):1-11. [doi: [10.1007/s41870-017-0080-1](https://doi.org/10.1007/s41870-017-0080-1)]
45. Breiman L. 랜덤 포레스트. In: *머신 러닝*. Vol 45. Springer; 2001:5-32. [doi: [10.1023/A:1010933404324](https://doi.org/10.1023/A:1010933404324)]

46. Cutler A, Cutler DR, Stevens JR. 랜덤 포레스트. In: 앙상블 머신 러닝: 방법 및 응용. Springer; 2012:157-175. [doi: [10.1007/978-1-4419-9326-7_5](https://doi.org/10.1007/978-1-4419-9326-7_5)]
47. Chen T, Guestrin C. Xgboost: 확장 가능한 트리 부스팅 시스템. arXiv. 2016년 3월 19일 온라인 게시된 사전 공개 논문. [doi: [10.48550/arXiv.1603.02754](https://doi.org/10.48550/arXiv.1603.02754)]
48. Rahman RU, Ahmad I, Li Z, Sparks R, Saad AB, Mullen A. Singletrome: 분석 및 향상을 위한 방법
단일 세포 분석을 위한 긴 비코딩 RNA를 이용한 전자체 분석. bioRxiv. 2022년 11월 2일. [doi: [10.1101/2022.10.31.514182v1](https://doi.org/10.1101/2022.10.31.514182v1)]
49. Goodfellow I, Bengio Y, Courville A. Deep Learning. The MIT Press; 2016. ISBN: 9780262035613 50. Misgar MM, Bhatia MPS. UMAP 특징을 이용한 iomt 시계열 데이터로부터 우울증 탐지. 2022년 국제 컴퓨팅, 통신 및 지능형 시스템 컨퍼런스(ICCCIS) 발표. 623-628; 2022.[doi: [10.1109/ICCCIS56430.2022.10037751](https://doi.org/10.1109/ICCCIS56430.2022.10037751)]
51. Ahmad I, A. Alqarni M, Ali Almazroi A, Tariq A. 기계 학습을 이용한 클락베이트 탐지의 실험적 평가
모델. 지능형 자동화 및 소프트 컴퓨팅. 2020;26(4):1335-1344. [doi: [10.32604/iasc.2020.013861](https://doi.org/10.32604/iasc.2020.013861)]
52. Muhammad D, Bendeache M. 블랙박스 해부: 설명 가능한 인공지능에 대한 체계적 검토
의료 영상 분석. 컴퓨터 구조 생명공학 저널. 2024년 12월;24:542-560. [doi: [10.1016/j.csbj.2024.08.005](https://doi.org/10.1016/j.csbj.2024.08.005)] [메드라인: [39252818](https://pubmed.ncbi.nlm.nih.gov/39252818/)]
53. Su-In L, Scott M. 모델 예측 해석을 위한 통합적 접근법. arXiv. 2017년 5월 22일 온라인 게시된 사전 출판 논문. [doi: [10.48550/arXiv.1705.07874](https://doi.org/10.48550/arXiv.1705.07874)]
54. Yang HJ, Cheng WJ, Hsiao MC 등. 주요 우울 장애 환자의 우울 증상 심각도 및 주의력과 관련된 휴식-활동 리듬: 12개월 추적 연구. Front Psychiatry. 2023;14:1214143. [doi: [10.3389/fpsyt.2023.1214143](https://doi.org/10.3389/fpsyt.2023.1214143)] [메드라인: [37663595](https://pubmed.ncbi.nlm.nih.gov/37663595/)]
55. Walker WH 2nd, Walton JC, DeVries AC, Nelson RJ. 일주기 리듬 장애와 정신 건강. Transl Psychiatry. 2020년 1월 23일;10(1):28. [doi: [10.1038/s41398-020-0694-0](https://doi.org/10.1038/s41398-020-0694-0)] [메드라인: [32066704](https://pubmed.ncbi.nlm.nih.gov/32066704/)]
56. Song YM, Jeong J, de Los Reyes AA 5th, et al. 주요 우울증 및 양극성 장애 환자의 수면, 일주기 리듬 및 기분 증상의 인과적 역할: 종단적 웨어러블 기기 데이터에서 얻은 통찰. EBioMedicine. 2024년 5월;103:105094. [doi: [10.1016/j.ebiom.2024.105094](https://doi.org/10.1016/j.ebiom.2024.105094)] [메드라인: [38579366](https://pubmed.ncbi.nlm.nih.gov/38579366/)]
57. Mimikou C, Kokkotis C, Tsiptisios D 외. 우울증 예측을 위한 설명 가능한 기계 학습. Diagnostics (바젤). 2025년 6월 2일;15(11):1412. [doi: [10.3390/diagnostics15111412](https://doi.org/10.3390/diagnostics15111412)] [메드라인: [40506984](https://pubmed.ncbi.nlm.nih.gov/40506984/)]
58. Salk RH, Hyde JS, Abramson LY. 대표적인 국가 표본에서 우울증의 성별 차이: 진단 및 증상의 메타 분석. Psychol Bull. 2017년 8월;143(8):783-822. [doi: [10.1037/bul0000102](https://doi.org/10.1037/bul0000102)] [메드라인: [28447828](https://pubmed.ncbi.nlm.nih.gov/28447828/)]
59. Maglione JE, Ancoli-Israel S, Peters KW, et al. 우울증 증상과 일주기 활동 리듬 장애
지역사회 거주 노인 여성. Am J Geriatr Psychiatry. 2014년 4월;22(4):349-361. [doi: [10.1016/j.jagp.2012.09.003](https://doi.org/10.1016/j.jagp.2012.09.003)] [메드라인: [23567424](https://pubmed.ncbi.nlm.nih.gov/23567424/)]
60. Wichers M, Groot PC, Psychosystems, ESM Group, EWS Group. 우울증에 대한 개인화된 조기 경고 신호로서의 임계 속도 저하. Psychother Psychosom. 2016;85(2):114-116. [doi: [10.1159/000441458](https://doi.org/10.1159/000441458)] [메드라인: [26821231](https://pubmed.ncbi.nlm.nih.gov/26821231/)]
61. Stiglic G, Kocbek P, Fijacko N, Zitnik M, Verbert K, Cilar L. 의료 분야에서 기계 학습 기반 예측 모델의 해석 가능성. WIREs Data Min & Knowl. 2020년 9월;10(5):e1379. [doi: [10.1002/widm.1379](https://doi.org/10.1002/widm.1379)]
62. Simula. 우울증 데이터셋. URL: <https://datasets.simula.no/depresjon/> [접근일: 2025년 9월 4일]

약어

ADASYN: 적응형 합성 샘플링 방법
 AI: 인공지능
 ANN: 인공신경망
 CNN: 합성곱 신경망
 DNN: 심층 신경망
 FN: 거짓 음성
 FP: 거짓양성
 LIME: 해석 가능한 모델에 구애받지 않는 설명
 LR: 로지스틱 회귀
 MADRS: 몽고메리-아스버그 우울증 평가 척도
 MCC: 매튜 상관 상수
 ML: 머신러닝
 MSE: 평균 제곱 오차
 NN: 신경망
 OvO: 1대1
 OvR: 1대 나머지

PSD: 전력 스펙트럼 밀도
RF: 랜덤 포레스트
SHAP: 셰플리 가산 설명
SMOTE: 합성 소수 집단 과표본 추출 기법
SVM: 서포트 벡터 머신
TN: 참음성
TP: 참양성
XAI: 설명 가능한 인공지능

존 토러스 편집; 이프라 샤힌, 비누 세리몬, 주무 저우 심사; 제출일: 2025년 2월 1일; 최종 수정본 접수일: 2025년 7월 13일; 게재 승인일: 2025년 7월 17일; 출판일: 2025년 9월 11일

인용 시 다음 형식을 사
용해 주십시오: Ahmed I, Brahmacharimayum A, Ali RH, Khan TA, Ahmad MO, 활동 데이터를
를 활용한 우울증 탐지 및 심각도 분류를 위한 설명 가능한 AI: 해석 가능한 프레임워크 개발 및 평가 연구, JMIR Ment Health 2025;12:e72038, URL: <https://mental.jmir.org/2025/1/e72038>
doi: [10.2196/72038](https://doi.org/10.2196/72038)

© Iftikhar Ahmed, Anushree Brahmacharimayum, Raja Hashim Ali, Talha Ali Khan, Muhammad Ovais Ahmad. 원래 JMIR Mental Health (<https://mental.jmir.org>)에 게시되었으며, 2025년 11월 9일, 이 논문은 크리에이티브 커먼즈 저작자표시 라이선스 (<https://creativecommons.org/licenses/by/4.0/>) 에 따
라 배포되는 오픈 액세스 논문입니다. JMIR Mental Health에 최초로 게재된 원본 저작물을 적절히 인용하는 조건으로, 모든 매체에서 제한 없이 사용, 배포 및 복제할 수 있습
니다. 전체 서지 정보 및 원본 출판물 링크(<https://mental.jmir.org/>)는 다음과 같습니다. 저작권 및 라이선스 정보도 반드시 포함되어야 합니다.