

논문 2023-18-08

머신러닝을 활용한 통계 분석 기반의 수면 호흡 장애 중증도 예측 (Severity Prediction of Sleep Respiratory Disease Based on Statistical Analysis Using Machine Learning)

김 준 수, 최 병 재*
(Jun-Su Kim, Byung-Jae Choi)

Abstract : Currently, polysomnography is essential to diagnose sleep-related breathing disorders. However, there are several disadvantages to polysomnography, such as the requirement for multiple sensors and a long reading time. In this paper, we propose a system for predicting the severity of sleep-related breathing disorders at home utilizing measurable elements in a wearable device. To predict severity, the variables were refined through a three-step variable selection process, and the refined variables were used as inputs into three machine-learning models. As a result of the study, random forest models showed excellent prediction performance throughout. The best performance of the model in terms of F1 scores for the three threshold criteria of 5, 15, and 30 classified as the AHI index was about 87.3%, 90.7%, and 90.8%, respectively, and the maximum performance of the model for the three threshold criteria classified as the RDI index was approx 79.8%, 90.2%, and 90.1%, respectively.

Keywords : Sleep-related Breathing Disorders, Machine Learning, Correlation Analysis, Prediction

1. 서 론

2019년 건강보험심사평가원이 발표한 자료에 따르면 수면 호흡 장애 (Sleep disordered breathing, SDB)를 겪은 사람의 통계는 2015년 약 45만 6,000명에서 5년간 연평균 8.1%씩 증가하여 2019년에는 약 63만 7,000명으로 연령대가 높아질수록 증가하는 추세를 보였다. 성인을 기준으로 한다면 한국 남성에서는 4.5%, 여성에서는 3.2% 정도로 흔한 질병이다 [1]. 수면 호흡 장애의 진단 및 평가는 임상증상만 가지고 판단하기에는 어려움이 있어 진단을 위해서는 수면다원검사 (Polysomnography, PSG)가 꼭 필요하다. 수면다원검사가 중요한 이유는 무호흡-저호흡 지수 (apnea-hypopnea index, AHI)나 호흡 장애지수 (respiratory disturbance index, RDI) 때문이다. 이 두 가지 지수는 수면 호흡 장애 진단에 필수적인 지수로 수면 중 호흡 상태를 객관적으로 보여주며, 수면다원검사를 통해서만 측정할 수 있다. 하지만, 수면다원검사는 다양한 센서를 부착하고 병원에서 수면을 취해야한다. 이는 환자가 병원이라는 익숙하지 않은 곳에서, 그림 1과같이 많은 검사 장치를 달고 검사실에서 수면을 취해야하기 때문에 수면이 평상시와 다를 수 있다. 또한, 수면다원검사 결과지를 해석하기 위해 오랜 시간이 소요되며, 많은 인력이 필요하다.

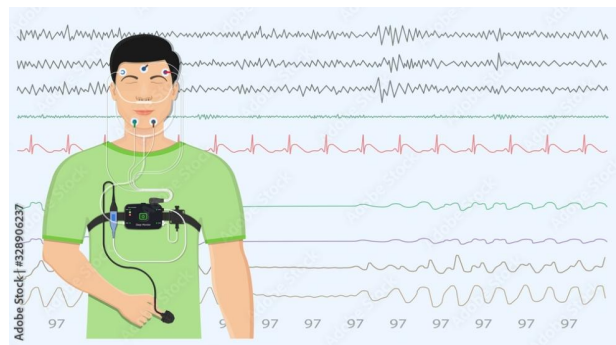


그림 1. 단계 선택법 과정에 따른 모델 설명력 변화
Fig. 1. Changes in model explanatory power according to the step-by-step selection process

수면다원검사를 통해 수면 호흡 장애를 진단하는 것이 중요한 이유는 여러 연구를 통해 보고되었다. Arzt 등의 연구에서 AHI (무호흡-저호흡 지수) > 20인 군이 AHI < 5인 군에 비해 훨씬 높은 뇌졸중 환자가 발생하였으며 [2], Yaggi 등의 연구에서는 AHI > 5인 군에서 뇌졸중의 발생이나 사망의 위험이 크다고 보고되었다 [3].

이러한 연구 결과들은 중등도 이상의 의미 있는 폐쇄성 수면무호흡일 때, 뇌졸중의 발생 위험도가 커진다는 것을 의미한다고 말한다 [4]. 또한, 고속도로 교통사고로 응급실에서 치료받은 운전자를 대상으로 한 조사 결과 무호흡-저호흡 지수가 10 이상일 때 6.3배로 교통사고의 위험이 증가하였다고 보고되었다 [5]. 이처럼, 수면 호흡 장애의 병 유무나

*Corresponding Author (bjchoi@daegu.ac.kr)
Received: Jan. 31, 2023, Revised: Feb. 22, 2023, Accepted: Apr. 13, 2023.
J. S. Kim: Daegu University (M.S.)
B. J. Choi: Daegu University (Prof.)
* 이 논문은 2023년도 정부 (산업통상자원부)의 재원으로 한국산업기술진흥원의 지원을 받아 수행된 연구임 (P0012724, 2023년 산업전문인력역량강화사업).

수면 호흡 장애의 중증도 정도에 따라 합병증의 발병 우려가 다르므로 수면 호흡 장애 검사를 통해 진단하여 중증도를 확실하게 아는 것이 중요하다.

현재 수면 호흡 장애를 진단하기 위한 다양한 연구가 진행되고 있다. Jeong은 상반신을 덮는 크기의 두께가 얇은 에어 매트리스를 이용하여 비접촉 방식으로 환자의 불편함을 최소화한 수면 호흡 장애 검출 시스템을 개발하였다. 연구한 알고리즘의 성능은 민감도 82.1%, 양성예측도 68.3%가 생성되었다 [6]. Cho 등의 연구에서는 평소 사용하는 스마트폰으로 수면 중 호흡음을 녹음하여 수면무호흡증을 예측하는 알고리즘을 개발하였다. 머신러닝 모델 중 Random Forest 알고리즘을 사용하여 무호흡-저호흡 지수 (AHI)로 구분된 임계값 5, 15, 30의 3가지 기준에 대해 이진 분류로 예측 모델을 생성하였다. 모델의 정확도는 3가지 임계값 기준에 대해 각각 88.2%, 82.3%, 81.7%가 나타났다 [7]. 두 연구 모두 병원에서 검사한다는 단점을 없애고, 검사 장치를 최소화하여 수면 호흡 장애를 검출하는 시스템을 연구하였다. 하지만, 연구 대상이 각각 9명과 423명으로 표본의 수가 작다는 단점이 있다.

본 연구에서는 AI 허브에서 제공하는 '수면 질 평가 및 수면장애 진단 이미지' 데이터를 활용하여 10,253명을 대상으로 연구하고자 한다. 연구는 최소한의 검사 장치를 가지고 가정에서 진단할 수 있는 시스템을 개발을 목적으로 한다. 중증도 예측에 AHI 지수와 RDI 지수를 각각 사용하여 예측 성능을 비교함으로써 폐쇄성 수면 무호흡증뿐만 아니라 다양한 무호흡증에 대해서도 중증도를 예측한다.

II. 예측을 위한 모델링

1. 데이터 수집

본 논문에서는 수면 호흡 장애 중증도 예측을 위한 연구로 사용된 데이터는 AI 허브 포털에서 제공하는 '수면 질 평가 및 수면장애 진단 이미지' 데이터를 활용한다. 원본 데이터는 IRB 승인하에 총 4개의 의료기관 (분당서울대병원, 한림대학교병원, 서울수면의원, 리엔홍이비인후과)에서 수집되었으며, 서울대학교병원에서 검증하였다. 해당 데이터는 양질의 데이터를 구축하고, 이를 이용한 인공지능 모델을 개발하여 자동화된 PSG 판독 시스템을 구축하기 위해 수집되었다. 이 데이터 셋의 구성으로는 호흡량, 호흡 노력, 코골이, 혈액의 산소포화도 정보 등의 신호를 30초 단위로 분류하여 나타낸 이미지와 이에 대한 Annotation 파일로 구성된다. 이 중 데이터는 이미지 데이터를 제외하고 Annotation 파일로 제공된 검사 결과 요약 데이터만을 사용하였다. 데이터 구축량은 검사 기준으로 약 10,000건으로 이는 2020년 기준, 해외에서도 가장 큰 규모의 PSG 데이터 셋이다. 세부 항목으로는 대상자의 특성을 나타내는 성별, 나이 BMI 지수가 있으며, 검사 결과로 취침 시간, 총 수면시간, REM 수면시간 등이 있다. 또한, 수면 시 센서 데이터로 측정된 AHI, RDI, RERA와 같은 지수 정보와 혈액의 산소포화도,

표 1. 데이터 셋 통계

Table 1. Data set Statistics

Classification	Type	Count	Total
Sex	Man	8,415	10,253
	Woman	1,828	
Age	under 20	1,467	10,253
	30's	2,474	
	40's	2,493	
	50's	2,069	
	over 60	1,723	
Test Type	PSG	7,745	10,253
	CPAP	2,149	
	MSLT	359	
Test Results	Embla	5,588	10,253
	Nox	4,665	

심박수 등의 항목이 있다.

구축된 데이터 셋의 통계는 표 1에서 확인할 수 있다. 단일 인종 (한국인)의 PSG 데이터로 이루어져 있으며, 수면무호흡증, 기면병, 불면증 환자의 검사 결과와 양압기의 적정 압력을 결정하기 위한 양압 적정 검사 결과를 포함한다. 남녀의 성비는 8,415:1,828로 이는 약 4:1 비율이다. 비율이 남성에 치우쳐 있는 것처럼 보이지만 이는 실제 환자의 분포를 반영하고 있다.

2. 수면 호흡 장애 진단기준

AHI 수면 호흡 장애의 지표는 시간당 무호흡 및 저호흡의 발생 횟수를 나타내는 무호흡-저호흡 지수 (apnea hyponea index, AHI)와 무호흡-저호흡 지수와 호흡 노력 관련 각성 (respiratory effort related arousal, RERA)의 발생 횟수를 포함하여 나타내는 호흡 장애지수 (respiratory disturbance index, RDI)를 흔히 사용한다. 무호흡 지수 (Apnea index, AI)와 저호흡 지수 (Hypopnea index, HI)를 합하여 무호흡-저호흡 지수라고 하고, 무호흡-저호흡 지수에 호흡 노력 관련 각성 지수 (RERA)를 합하여 호흡 장애 지수라고 한다.

현재 수면 호흡 장애를 진단하는데 AHI 지수를 사용할 것인가 또는 RDI 지수를 사용할 것인가도 통일되지 않았다. 실제로 AHI 지수가 현재 수면무호흡증 진단 및 중증도 평가에 주로 사용되고 있지만, 중증도를 판단하기에는 충분하지 않을 수 있다고 말한다. 이에 본 연구에서는 수면 호흡 장애 중증도를 진단하기 위해 호흡 장애지수 (RDI)를 함께 고려하여 수면 호흡 장애 중증도를 예측한다. 호흡장애 지수는 AHI 지수와 RERA 지수가 함께 포함되어 있어 폐쇄성 수면무호흡과 저호흡뿐 아니라 중추성 무호흡과 같은 다양한 무호흡증을 진단할 수 있다. 수면 호흡 장애의 중증도

표 2. 수면 호흡 장애 진단기준

Table 2. Diagnostic Criteria for Sleep Respiratory Disorder

SDB severity	AHI & RDI
Normal	< 5
Mild	5 ~ 15
Moderate	15 ~ 30
Severe	30 <

는 AHI 지수나 여기에 RERA 지수를 포함하여 RDI 지수를 계산하며 표 2와 같이 총 4단계로 분류한다. AHI 지수나 RDI 지수가 5 미만일 경우 정상, 5에서 15 미만일 경우 경증, 15에서 30일 경우 중등도, 30을 초과할 경우 중증으로 정의한다 [8].

3. 예측에 사용된 모델

1) 랜덤 포레스트

랜덤 포레스트는 여러 개의 모델을 결합해 하나의 모델보다 더 우수한 예측 성능을 갖는 앙상블 (Ensemble) 계열의 모델이다. 랜덤 포레스트는 의사결정나무 (Decision Tree)가 확장된 형태로서 의사결정나무가 가진 훈련 데이터에 대한 성능과 테스트 데이터에 대한 성능 차이가 커지는 과대 적합 (Overfitting)과 낮은 정확도 등의 문제를 개선하기 위해서 개발되었다 [9]. 랜덤 포레스트는 모형별로 변수 선택 및 표본 선정에 있어 랜덤으로 부여하고 독립적인 의사결정나무를 반복적으로 만들어, 의사결정나무의 분산을 낮추고 낮은 편향 (bias)을 유지하여 예측 오차를 줄일 수 있다 [10]. 랜덤 포레스트에서 여러 개의 의사결정나무를 만들기 위해 관측치와 예측인자에 대한 랜덤 추출을 반복하며, 추출에는 의사결정나무와 같이 부트스트랩 (Boostrap) 방법을 사용한다 [11]. 랜덤 포레스트는 제일 먼저 배깅 (Bagging)을 거친다. 배깅은 Bootstrap Aggregating의 줄임말로, 모집단에서 추출된 training dataset에서 랜덤 복원 추출한 여러 개의 부트스트랩 자료를 생성하여 각각의 예측 모형을 생성한 후 조합하여 최종 예측 모형을 만든다 [12]. 의사결정나무의 각 마디에서 독립변수를 선택할 때도 무작위성을 더해 각 의사결정나무 모형 사이의 상관성을 감소시켜 과적합 문제에서도 자유로우며 예측력을 높일 수 있다. 또한, 모든 변수들을 균형 있게 학습할 수 있으므로 학습과 판별을 빠르게 처리하고, 학습 데이터의 노이즈에도 강하다는 장점이 있다 [13].

2) AdaBoost

Adaboost는 Adaptive boost의 줄임말로 boosting은 다수의 분류기에서 예측력이 상대적으로 낮은 분류기들을 결합하여 예측력이 높은 분류기로 바꾸는 알고리즘을 이용하여 전체 모델 내의 분산을 줄이고, 예측력을 높이는 방법이다 [14]. Adaboost는 예측 성능이 낮은 약 학습기가 여러 단계를 거치면서 이전 단계에서 오 분류로 판단된 학습 데이터

에 대해 가중치를 부여한다. 그 후 다음 학습기에 가중치를 적용하여 학습 데이터에 영향을 받도록 하여 단계별 선형 분류가 다르게 나타나도록 반복 학습하는 알고리즘이다 [15]. 하나의 가지가 잘못 분류한 sample에 대해서는 다음 가지로 넘겨줄 때 가중치를 더 높여서 넘겨준다. 이후 단계에서는 이전 가지에서 잘못 분류한 sample에는 sample weight를 높여주고, 이전 가지에서 제대로 분류한 sample에는 sample weight를 줄여준다. 이러한 방법으로 이전 모델에서 오 분류된 특징에 대해 높은 가중치를 부여하고 정 분류된 특징에 대해서는 낮은 가중치를 부여하여 여러 약한 학습기들에 대해 순차적으로 진행하면서 학습 데이터의 가중치를 재조정하며 학습한다. 약한 학습기를 순차적으로 하나씩 학습할 때 이전의 학습기에서 오 분류된 결과를 다음 학습기가 학습할 때 사용하여 이전 학습기의 단점을 보완한다 [16].

3) XGBoost

XGBoost는 tree 기반 모델이나 선형 모델에서의 과적합 문제를 해결하고, 방대한 데이터 셋의 안정성을 높이면서 훈련 속도를 향상시키기 위해 개발된 방법이다. Classification and Regression Tree (CART) 알고리즘의 기반으로 Boosting 기법을 활용한 의사결정나무 기반의 앙상블 모델이며, 그림 2와 같이 새로운 데이터에 이전 학습 모델의 결과를 반영하여 성능을 더욱 향상시키는 방법으로 연속적인 훈련 알고리즘의 특징을 가지고 있다 [17].

XGBoost는 Gradient Boosting과 비슷하지만, 학습 데이터에 가중치를 적용하는 방법에 차이가 있다. Gradient Boosting에서는 tree pruning에서 negative loss가 발생하면 그 과정을 멈추게 된다. 하지만, XGBoost는 모델을 수행할 때 파라미터로 지정한 max_depth까지 진행하고 난 후 결과를 비교하여 loss function에서 개선이 일정 수준에 도달하지 못하게 된다면 역방향으로 pruning을 진행한다 [18, 19]. XGBoost에서는 임계점이 되는 지점을 Quantiles로 나눠서 후보 (Candidate)를 선정하는 방식으로 부스팅의 느린 학습 속도를 개선한다 [20].

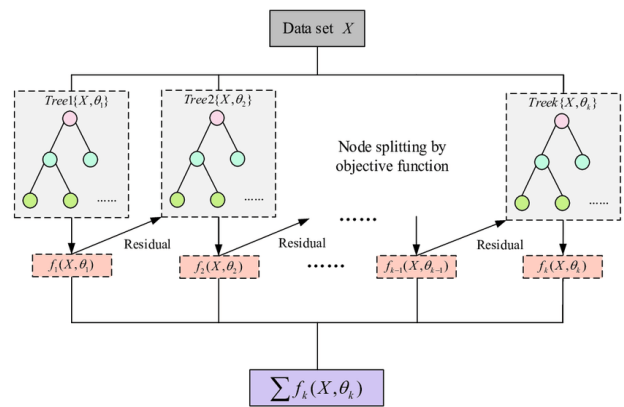


그림 2. XGBoost 흐름도 [21]
Fig. 2. Flow chart of XGBoost

표 3. AHI 지수 임계값 5 기준 상관분석 결과

Table 3. Results of correlation analysis based on AHI index critical point 5

	Variable	Correlation	p-value
1	Sex	-0.15	0.00
2	Age	0.16	0.00
3	BMI	0.07	0.00
4	Time in Bed (TIB)	0.04	0.00
5	Total Sleep Time (TST)	-0.07	0.00
6	Sleep Latency	0.00	0.99
...			
26	Lowest Oxygen Desaturation	-0.37	0.00
27	Mean Heart Rate	0.10	0.00
28	Lowest Heart Rate	0.21	0.00

III. 시뮬레이션 및 결과

1. 변수 선택과정

본 논문에서는 수면 호흡 장애 중증도를 예측하기 위해 '수면 질 평가 및 수면장애 진단 이미지'에서 제공하는 약 80개의 변수를 3단계에 걸쳐 정제하였다. 통계적인 변수 선택 방법인 두 번째와 세 번째 단계는 목표변수에 따라 상관성과 유의한 변수가 달라지므로 매번 새롭게 진행하였다. 아래 변수 선정 과정에서는 AHI 지수 임계값 5를 기준으로 하여, 이진 분류를 수행했을 때를 예시로 보여준다.

1) 변수 선정 방법

첫 번째 변수 선택과정은 가정에서 간단하게 측정 가능한 변수를 임의로 선택하였다. 먼저, 개인의 특성을 나타낼 수 있는 성별, 나이, BMI 지수를 선택하였다. 가정에서 측정 가능한 변수를 선택한 기준은 웨어러블 기기인 갤럭시 워치 4와 애플 워치 8에서 제공하고 있는 수면 분석 애플리케이션에서 제공되는 항목을 사용하였다. 수면 분석 애플리케이션에서 제공되고 데이터는 총 수면시간, 수면 단계별 시간, 혈중 산소 포화도, 심박수 등이 있다. 총 수면시간에서 취침 시간과 깨어있는 시간을 추가로 사용하였고, 수면시간 정보는 개인에 따라 총 수면 시간이 다르므로 단순 수면시간 정보만을 사용하는 것이 아닌 총 수면 시간 중 수면 단계별 시간 비율을 함께 사용하였다. 혈중 산소 포화도 정보를 통해 얻을 수 있는 평균 산소 포화도, 최저 산소 불포화도, 90% 이하의 산소포화율 변수를 사용하였다. 심박수 정보로는 평균 심박수, 최고 심박수, 최저 심박수 데이터를 사용하였다. 이 변수 선택과정으로 약 80개의 제공된 변수에서 29개가 선택되었다. 또한, 데이터는 수집하는 과정에서 학습에

표 4. 단계 선택법 후 AHI 지수 임계값 5 기준 유의한 변수

Table 4. Significant variables according to the AHI index threshold of 5 after stepwise selection

	Variable
1	Ratio of Snoring Time to TST
2	3% Oxygen Desaturation Index
3	Lowest Oxygen Desaturation
4	Ratio of Saturation below 90%
5	Mean Oxygen Saturation
6	Lowest Heart Rate
7	Stage 1 Time
8	Ratio of Sleep Time in the Lateral Position to TST
9	Ratio of Sleep Time in the Supine Position to TST
10	Stage 3 Time
11	Ratio of Stage 3 Time to TIB

용이하도록 하나로 취합하고 형식을 통일하는 것이 필요하므로 변수 정제 과정에서 데이터 형식도 통일하였다 [22].

2) 상관분석

두 번째 변수 선택과정은 상관분석이다. 상관관계 분석은 통계학에서 주로 사용하는 방법으로 두 변수 간의 상관계수(Correlation coefficient)를 구하여 어떤 선형적 관계가 있는지에 대해 분석하는 방법이다. 이진 분류를 위해 만들어진 목표변수와 첫 번째 단계에서 선택된 29개 변수와의 상관분석을 진행한다. 상관관계 결과는 표 3에서 확인할 수 있다.

그 결과 수면 대기시간 (Sleep Latency) 변수와 최고 심박수 (Highest Heart Rate) 변수의 p-value 값이 각각 0.99과 0.50로 유의수준 0.05보다 크므로 유의하지 않아 제거한다. 또한, 상관계수의 크기가 너무 작은 경우 상관성이 낮다고 해석할 수 있으므로 $-0.2 \sim 0.2$ 의 상관계수를 가진 변수들은 제거한다. 분석 결과 p-value 값이 크거나 상관계수의 크기가 작은 경우를 제외하여 29개 변수 중 12개가 선택되었다.

3) 변수 선정 방법

마지막 변수 선택과정은 변수 선택법 중 단계 선택법을 사용하였다. 단계 선택법은 전진 선택법에서 후진 제거법을 추가한 방법이다. 먼저, 전진 선택법과 같이 유의한 변수를 추가한 다음 새로 추가된 변수에 대해 모델의 설명력이 향상하는지 감소하는지 판단한다. 설명력이 향상하는 변수의 경우 모델에 추가하고, 설명력이 감소하는 변수의 경우 모델에서 제외하는 방법으로 변수를 선택한다.

단계 선택법을 통해 최종적으로 선택된 변수는 표 4와 같다. 기존 상관분석에서 유의했던 변수 12개 중 취침 시간에

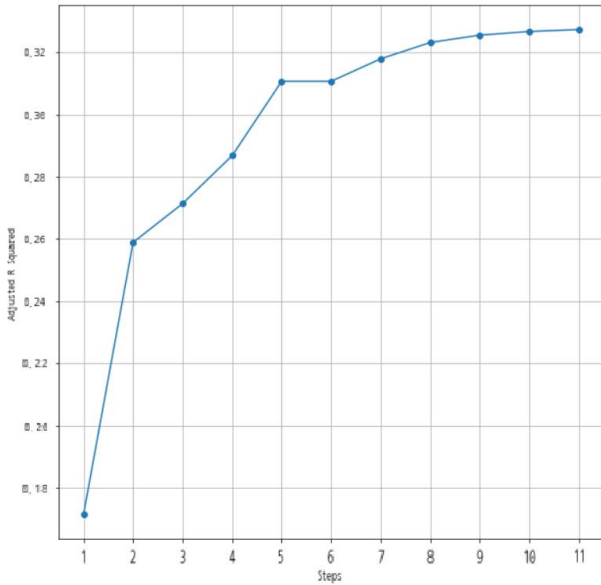


그림 3. 단계 선택법 과정에 따른 모델 설명력 변화
 Fig. 3. Changes in model explanatory power according to the step-by-step selection process

대한 1단계 시간의 비율 (Ratio of Stage 1 Time to TIB) 변수를 제외하고 11개가 선택되었다.

변수 선택에 따른 모델의 설명력의 그래프 변화는 그림 3과 같다. 단계 선택법에서는 변수를 추가하였을 때 설명력이 감소하면 그 변수를 삭제한다. 이로 인해 그래프에서 변수 추가에 따른 설명력 감소는 없으며, 변수가 추가될 때마다 설명력이 같거나 커지는 것을 확인할 수 있다.

이 예시의 변수 목록을 제외하고, 다른 기준에서의 중증도 예측에서도 7~13개의 변수가 사용된다. 대부분 '수면시간 중 코골이 시간 비율', '3% 산소 불포화 지수', '평균 산소포화도', '1단계 렘 수면시간' 변수가 포함되고, '90% 이하 산소포화율', '3단계 렘 수면시간' 변수도 6번의 이진 분류 중 4번 포함되었다.

2. 중증도 예측 결과

본 논문에서는 AHI 지수와 RDI 지수의 임계값 5, 15, 30에 따라 3가지 임계값 기준에 대해 이진 분류를 각각 진행하였다. 아래 결과에서는 랜덤 포레스트, AdaBoost, XGBoost 3가지 모델을 이용한 수면 호흡 장애 중증도 예측 결과를 비교한다.

1) AHI 지수 기준 중증도 예측 결과

AHI 지수를 기준으로 한 중증도 예측 결과는 표 5에서 확인할 수 있다, 임계값 5를 기준으로 한 이진 분류에서 가장 좋은 예측 성능을 보여준 모델은 랜덤 포레스트 모델로 F1 Score 값은 0.8727이 생성되었다. 임계값 15의 기준에서는 랜덤 포레스트 모델의 F1 Score 값이 0.9067, 임계값 30 기준에서는 0.9078로 F1 Score가 생성되었다. 세 가지 이진 분류에서 랜덤 포레스트 모델이 가장 예측 성능이 좋았으며, 그다음으로 XGBoost, AdaBoost 모델 순으로 높은 예측 성

표 5. AHI 지수 기준 중증도 예측 결과 비교
 Table 5. Comparison of severity prediction results based on AHI index

	AHI Critical Point		
	5	15	30
Random Forest	0.8727	0.9067	0.9078
AdaBoost	0.8616	0.8952	0.8953
XGBoost	0.8533	0.9010	0.9001

(Index : F1 score)

표 6. RDI 지수 기준 중증도 예측 결과 비교
 Table 6. Comparison of severity prediction results based on RDI index

	RDI Critical Point		
	5	15	30
Random Forest	0.7983	0.9019	0.9007
AdaBoost	0.7737	0.8956	0.8964
XGBoost	0.7840	0.9010	0.9107

(Index : F1 score)

능을 보여주었다. 임계값이 높을수록 중증을 의미하므로, 경증에서의 분류 성능보다 중증도와 중증의 분류 성능이 우수하다고 할 수 있다.

2) RDI 지수 기준 중증도 예측 결과

RDI 지수를 기준으로 한 중증도 예측 결과를 표 6에서 확인할 수 있다, 임계값 5를 기준으로 한 이진 분류에서 가장 좋은 예측 성능을 보여준 모델은 랜덤 포레스트 모델로 F1 Score 값은 0.7983이 생성되었다. 임계값 15의 기준에서는 랜덤 포레스트 모델의 F1 Score 값이 0.9019로 XGBoost 0.9010과 큰 차이가 없는 것을 볼 수 있다. 임계값 30 기준에서는 랜덤 포레스트 0.9007보다 XGBoost 0.9107의 성능이 더 우수하다. 세 가지 이진 분류에서 랜덤 포레스트 모델의 예측 성능이 대부분 좋은 예측 결과를 보였고, 30 기준에서는 XGBoost 모델의 예측 성능이 가장 높았다.

IV. 결론

본 연구에서는 AI 허브에서 제공하는 '수면 질 평가 및 수면장애 진단 이미지' 데이터를 활용하여 머신러닝을 활용한 통계 분석 기반의 수면 호흡 장애 중증도 예측 연구를 진행하였다. 수면 호흡 장애는 매년 증가하는 추세를 보이고 있으며, 한국 남성에서는 4.5%, 여성에서는 3.2% 정도로 흔한 질병이다. 주위에서 흔히 찾아볼 수 있는 코골이와 수면무호흡증은 어느 정도 무호흡증이 치료가 필요한 상황인지 판단이 필요하다. 수면 호흡 장애를 진단하기 위해서는 현재 수면다원검사가 필수적으로 필요하다. 수면 호흡 장애 진단에 필수적이면서 수면 중 호흡 상태를 객관적으로 보여

주는 AHI 지수와 RDI 지수를 수면다원검사를 통해서만 측정할 수 있기 때문이다. 하지만, 수면다원검사는 여러 가지 단점이 있어 이러한 수면다원검사를 대체하기 위해 다양한 연구가 진행되고 있다.

본 논문에서는 서울대학교병원에서 우리나라 환자를 대상으로 구축한 '수면 질 평가 및 수면장애 진단 이미지' 데이터를 활용하였다. 주어진 데이터에서 수면 호흡 장애 중증도 예측을 위한 간단하고 정확도 높은 모델을 만들기 위해 3단계의 변수 선택과정을 사용하였다. 첫 번째로 웨어러블 기기에서 측정 가능한 항목을 임의로 선택하고, 두 번째로 상관분석을 진행하였다. 상관분석을 진행한 후 목표변수와의 상관성이 유의한 변수들만을 선택하여 마지막으로 변수 선택법 중 단계 선택법을 이용하여 변수를 정제하였다. 최종적으로 선택된 변수들을 이용하여 3가지 머신러닝 모델에 활용하였다. 여기에서 활용한 3가지 모델은 랜덤 포레스트, AdaBoost, XGBoost 모델이다. 3가지 모델을 이용하여 AHI 지수와 RDI 지수별 임계값에 따른 중증도를 예측하였다. 예측 결과는 3단계 변수 선택과정을 통해 정제되어 최종으로 선택된 변수 목록을 보였으며, F1 Score 성능 지표로 모델 성능을 평가하였다. 예측 결과 대부분 랜덤 포레스트 모델이 3가지 모델 중 가장 높은 예측 성능을 보여주었다. AHI 지수로 구분된 5, 15, 30 임계값 기준에 대한 모델의 최대 성능은 각각 약 87.3%, 90.7%, 90.8%가 나타났다. RDI 지수로 구분된 5, 15, 30 임계값 기준의 모델 성능은 각각 약 79.8%, 90.2%, 90.1%로 나타났다. 예측 성능만 비교하면, AHI 지수를 사용한 결과에 비해 조금 낮은 예측 성능을 보여주었다. 하지만, AHI 지수는 수면무호흡증의 중증도만 진단하는 반면, RDI 지수는 호흡 노력 관련 각성 지수를 포함하고 있어 좀 더 넓은 범위의 수면 호흡 장애를 진단할 수 있다.

추후 분석에서는 조금 더 추가된 센서 데이터를 사용하더라도 수면다원검사를 진행하지 않고 더 정확한 수면 호흡 장애의 중증도를 예측하는 알고리즘을 연구해보려고 한다. 수면 호흡 장애를 더욱 정확하게 예측 가능하다면, 자신의 수면 호흡 장애를 빨리 알아차리고 중증도가 심해지기 전 예방을 할 수 있을 것이다. 또한, 매년 증가하고 있는 수면 호흡 장애 환자로 인한 수면 검사실 부족과 인력 부족을 해결할 수 있을 것으로 기대된다.

References

- [1] J. K. Kim, K. H. In, J. H. Kim, S. H. You, K. H. Kang, J. J. Shim, S. Y. Lee, J. B. Lee, S. G. Lee, C. Park, C. Shin, "Prevalence of Sleep-disordered Breathing in Middle-aged Korean Men and Women," *American Journal of Respiratory and Critical Care Medicine*, Vol. 170, No. 10, pp. 1108 - 1113, 2004.
- [2] M. Arzt, T. Young, L. Finn, J. B. Skatrud, T. D. Bradley, "Association of Sleep-disordered Breathing and the Occurrence of Stroke," *American Journal of Respiratory and Critical Care Medicine*, Vol. 172, No. 11, pp. 1447-1451, 2005.
- [3] H. K. Yaggi, J. Concato, W. N. Kernan, J. H. Lichtman, L. M. Brass, V. Mohsenin, "Obstructive Sleep Apnea as a Risk Factor for Stroke and Death," *New England Journal of Medicine*, Vol. 353, No. 19, pp. 2034-2041, 2005.
- [4] S. J. Choi, "Complications of Obstructive Sleep Apnea," *The Korean Journal of Medicine*, Vol. 89, No. 1, pp. 13-20, 2015 (in Korean).
- [5] J. Teran, A. Jimenez, J. Cordero, Cooperative Group Burgos - Santander, "The Association Between Sleep Apnea and the Risk of Traffic Accidents," *New England Journal of Medicine*, Vol. 340, No. 11, pp. 847-851, 1999.
- [6] P. S. Jeong, "Development of Non-contact Sleep-disordered Breathing Detection System using Air-Mattress," Master Degree, Yonsei University, 2016 (in Korean).
- [7] P. S. Jeong, "Development of Non-contact Sleep-disordered Breathing Detection System using Air-Mattress," Master Degree, Yonsei University, 2016 (in Korean).
- [8] S. F. Quan, J. C. Gillin, M. R. Littner, J. W. Shepard, "Sleep-related Breathing Disorders in Adults: Recommendations for Syndrome Definition and Measurement Techniques in Clinical Research," *Editorials. Sleep (New York, NY)*, Vol. 22, No. 5, pp. 662-689, 1999.
- [9] R. M. Kim, K. M. Kim, J. H. Ahn, "Comparison Between Random Forest and Recurrent Neural Network for Photovoltaic Power Forecasting," *Journal of Korean Society Of Environmental Engineers*, Vol. 43, No. 5, pp. 347-355, 2021 (in Korean).
- [10] K. H. Hong, "A Predictive Model for Suicidal Ideation of Adolescents Using Random Forests Machine Learning Algorithm," *Korean Journal of Social Welfare*. Vol. 72, No. 3, pp. 157-180, 2020.
- [11] B. S. Seo, "A Study on the Prediction Model of the Elderly Depression Based on Random Forest," Master Degree, Dankook University. 2019 (in Korean).
- [12] B. S. Seo, "A Study on Prediction of Excessive Outpatient Care Users Using Random Forest and Artificial Neural Network," Master Degree, Namseoul University, 2019 (in Korean).
- [13] Y. J. Oh, H. C. Jung, "Development of Galaxy Image Classification Based on Hand-crafted Features and Machine Learning," *IEMEK Journal of Embedded Systems and Applications*, Vol. 16, No. 1, pp. 17-27, 2021 (in Korean).
- [14] J. H. Han, D. K. Ko, H. J. Choe, "Predicting and Analyzing Factors Affecting Financial Stress of Household using Machine Learning: Application of XGBoost," *Korean Society of Consumer Studies*, Vol. 30, No. 2, pp. 21-43, 2019 (in Korean).
- [15] J. H. Kim, "Hybrid AdaBoost for Multiuser Priority Analysis," Master Degree, Chung-Ang University, 2016 (in Korean).

[16] J. H. Kim, "A Comparative Analysis of Intelligent Classification Algorithms for Early Diabetes Prediction," Master Degree, Chonnam National University, 2021 (in Korean).

[17] J. H. Kim, "Comparative Analysis of Corporate Distress Predictive Models Using XGBoost," Master Degree, Dongguk University, 2019 (in Korean).

[18] T. Chen, C. Guestrin, "Xgboost: A Scalable Tree Boosting System," In Proceedings of the 22nd acm Sigkdd International Conference on Knowledge Discovery and Data Mining. pp. 785-794, 2016.

[19] J. H. Friedman, "Greedy Function Approximation: a Gradient Boosting Machine," Annals of Statistics. pp. 1189-1232, 2001.

[20] C. J. Kim, "Price Prediction on Building Prices with XGBoost Algorithm : Focusing on Eco-friendly Certification System," Master Degree, Sungkyunkwan University. 2022 (in Korean).

[21] Degradation State Recognition of Piston Pump Based on ICEEMDAN and XGBoost - Scientific Figure on ResearchGate. Available from: https://www.researchgate.net/figure/Flow-chart-of-XGBoost_fig3_345327934 [accessed 24 Mar, 2023]

[22] D. G. Kim, Y. S. Park, L. J. Park, T. Y. Chung, "Developing of New a Tensorflow Tutorial Model on Machine Learning : Focusing on the Kaggle Titanic Dataset," IEMEK Journal of Embedded Systems and Applications, Vol. 14, No. 4 pp. 207-218, 2019 (in Korean).

Jun-Su Kim (김 준 수)



2021 Department of Statistics and Big Data, Daegu University (B.S.)
 2023 Department of IT Convergence Engineering, Daegu University (M.S.)

Field of Interests: Statistics & Big Data, Machine Learning & Artificial Intelligence Theory
 Email: ingkkings@naver.com

Byung-Jae Choi (최 병 재)



1987 Department of Electric Engineering, Kyungpook National University (B.S.)
 1989 Department of Nuclear Engineering, KAIST (M.S.)
 1998 Department of Electrical and Electronic Engineering, KAIST (Ph.D.)
 1999.03~School of Electronic and Electrical

Engineering at Daegu University (Prof.)
 Field of Interests: Intelligent Controls & Systems, Artificial Intelligence Theory & Applications
 Email: bjchoi@daegu.ac.kr