



# Diagnostic Accuracy of Different Machine Learning Algorithms for Obstructive Sleep Apnea

Hyun-Woo Kim<sup>1</sup>, Euihwan Park<sup>2</sup>, Dae Jin Kim<sup>3</sup>, Sue Jean Mun<sup>4</sup>,  
Jiyoung Kim<sup>5</sup>, Gha-Hyun Lee<sup>5</sup>, Jae Wook Cho<sup>1</sup>

Departments of <sup>1</sup>Neurology and <sup>4</sup>Otorhinolaryngology, Pusan National University Yangsan Hospital, Yangsan,

<sup>2</sup>Department of Economics, Hannam University, Daejeon,

<sup>3</sup>Department of Biomedical Laboratory Science, Kyungbuk University, Porcheon,

<sup>5</sup>Department of Neurology, Pusan National University Hospital, Busan, Korea

## 다양한 머신 러닝 알고리즘을 이용한 폐쇄수면무호흡 진단의 정확도

김현우<sup>1</sup>, 박의환<sup>2</sup>, 김대진<sup>3</sup>, 문수진<sup>4</sup>, 김지영<sup>5</sup>, 이가현<sup>5</sup>, 조재욱<sup>1</sup>

양산부산대학교병원 신경과학교실, <sup>1</sup>이비인후과학교실, <sup>4</sup>한남대학교 경상대학 경제학과, <sup>2</sup>

경북대학교 임상병리학과, <sup>3</sup>부산대학교병원 신경과학교실<sup>5</sup>

Received October 22, 2020  
Revised November 12, 2020  
Accepted November 17, 2020

### Address for correspondence

Jae Wook Cho, MD, PhD  
Department of Neurology,  
Pusan National University  
Yangsan Hospital,  
20 Geumo-ro, Mulgeum-eup,  
Yangsan 50612, Korea  
Tel: +82-55-360-2122  
Fax: +82-55-360-2152  
E-mail: sleepcho@pusan.ac.kr

**Objectives:** The objective of this study was to develop models for predicting obstructive sleep apnea (OSA) based on easily obtainable clinical information of patients using various machine learning techniques.

**Methods:** We used a data set that included the records of 1,368 patients, in which 1,074 patients were male (78.5%), and 294 patients were female (21.5%). We randomly divided the data into a training set (1,000) and test set (368). Five machine learning methods, i.e., support vector machine model, lasso logit model, naïve bayes, discriminant analysis, and K-nearest neighbor (KNN), with a 10-cross fold technique were used with the proposed model to predict OSA. We evaluated the accuracy, sensitivity, specificity, and precision of each model for three thresholds [Apnea-Hypopnea Index (AHI)≥5, AHI≥15, and AHI≥30]. **Results:** Among the machine learning techniques, KNN showed the best results compared to the other techniques. The accuracy, sensitivity, specificity, and precision of OSA prediction were 87.0%, 91.0%, 74.4%, and 91.9%, respectively, based on AHI≥5. When the threshold of OSA was AHI≥15 or AHI≥30, KNN provided lower accuracy (79.6% each) and precision (79.0% and 68.7%), which were still higher than those of the other techniques. **Conclusions:** The model derived from the KNN technique exhibited the best performance based on its highest level of accuracy. We demonstrate that this model is a useful tool for predicting OSA.

J Sleep Med 2020;17(2):128-137

**Key Words:** Obstructive sleep apnea, Apnea, Machine learning, Algorithm.

## 서 론

폐쇄수면무호흡(obstructive sleep apnea, OSA)은 비교적 흔한 질환으로 30세 이상의 성인에서 5~14%의 유병률을 가지는 것으로 알려져 있다.<sup>1-3</sup> 국내에서도 19세 이상 성인의 15.8%(남성 19.8%, 여성 11.6%)가 가지고 있는 것으로 보고

되었다.<sup>4</sup> OSA는 고혈압, 당뇨, 심뇌혈관질환뿐만 아니라, 인지기능 장애 및 우울증 등 다양한 질환의 위험인자로 알려져 있어 적극적인 관리와 치료가 필요하다.<sup>5-7</sup> 하지만 여러 가지 이유로 높은 유병률에 비해 OSA로 확진받는 진단율은 낮은 편이다. 코골이나 수면무호흡을 숙면을 취하는 모습이라 생각하거나, 반대로 수면무호흡에 대한 부정적인 인식 때문에 진단을 꺼리는 경우도 많아 수면클리닉을 방문하지 않는 경우도 있다.<sup>8,9</sup> 또한, 환자 본인이 코골이나 수면무호흡을 호소하거나 같이 잠을 자는 보호자들의 증언에 의존하여 수면다

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<https://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.

원검사를 진행하는 경우가 많은데, 1인 가구가 늘면서 같이 사는 사람이 없거나 환자 자신이 증상을 인지하지 못하는 경우도 많아 OSA를 진단을 받지 못하는 사람이 늘어날 가능성이 있다. 따라서 환자나 보호자가 코골이나 수면무호흡을 호소하지 않아도 OSA가 예상되는 사람을 조기에 선별하고 이들이 수면클리닉을 방문하여 진단을 받도록 할 수 있다면 OSA의 치료와 환자의 삶의 질 향상에 도움이 될 것이다.

OSA의 선별검사를 위해 베를린 설문, sleep apnea clinical score, STOP, STOP-Bang 설문 등 증상과 징후를 토대로 진단할 수 있도록 점수화된 설문 도구들이 사용되고 있다. 이들 설문들은 편리하게 사용할 수 있다는 장점이 있으나 설문 작성자의 주관에 의존할 수밖에 없으며, 주로 서양인들을 기반으로 하여 만들어진 도구로 동양인에 바로 적용하기 위해서는 유효성 검증을 거쳐야 하는 단점이 있다.<sup>10,11</sup>

머신 러닝(machine learning)은 인공지능을 구현할 수 있는 한 분야로서 대용량의 데이터를 구체적인 알고리즘을 적용하여 컴퓨터를 통해 학습시키고 분석 작업을 수행할 수 있도록 기술을 개발하는 과정이다.<sup>12</sup> 각 변수들의 상호작용을 통계적으로 파악하여 규칙을 찾아내고 이를 기반으로 하여 직접 프로그래밍하여 공식을 개발하는 방식이 아니고, 다양한 예시 자료들을 제공하여 학습을 시키고 이를 통해서 결론을 도출할 수 있는 알고리즘을 만들어낸다는 점에서 기존의 예측 모델과 차별성이 있다. 수집된 데이터를 이용하여 높은 정확도와 적합도를 제시할 수 있는 방향으로 학습되고, 새로운 자료가 추가로 입력되었을 때 이를 통해 재학습하여 알고리즘을 성장시키는 방식이다.<sup>13</sup> 최근 의료 연구의 분야에서 머신 러닝 기법이 많이 적용되고 있는데,<sup>14,15</sup> OSA는 다른 질병에 비해 질병의 이환에 영향을 미치는 다양한 변수들이 존재하며 이들 간의 복잡한 상관관계가 있어 머신 러닝 기법을 적용하여 예측 모델을 제시하기 적합할 것으로 기대된다.

본 연구에서는 환자에게서 쉽게 얻을 수 있는 정보와 간단한 신체 계측만을 이용하여 OSA를 진단할 수 있는 머신 러닝 모델을 개발하고자 하였다. 다양한 머신 러닝 알고리즘을 적용하여 실제 수면다원검사를 통해 측정한 Apnea-Hypopnea Index(AHI)와 비교하여 정확도를 산출했고, OSA 중증도에 따라 각 알고리즘 모델마다 어느 정도의 효용성을 나타나는지 검증해 보았다.

## 방 법

### 대 상

본 연구는 2012년 7월~2019년 5월까지 양산부산대학교병원 수면클리닉에서 19세 이상 성인에게 시행한 수면다원검

사를 분석하였다. 총 수면 시간이 4시간 이하인 경우, 측정치에 결측값이 있는 경우, 분할 수면다원검사(split-night polysomnography) 및 양압기 사용을 위한 적정 압력(continuous positive airway pressure titration)을 위한 수면다원검사 자료는 제외하여 총 1,368건의 검사를 대상으로 하였다. 본 연구는 양산부산대학교병원 의학연구윤리 심의위원회(Institutional Review Board)의 승인하에 진행되었다(승인번호 05-2020-154).

### 신체 계측과 설문

환자들의 의무기록을 바탕으로 나이, 성별, 키, 몸무게, 체질량지수(body mass index, BMI)를 조사하였다. 신체 계측은 수면다원검사 직전에 시행하여 기록된 결과치를 이용하였다. 목 둘레는 앉은 자세와 누운 자세에서 각각 후두용기부의 상연에서 수평으로 둘레를 측정하였고, 허리 둘레는 늑골 하단과 장골능과의 중간 부위 둘레로 측정하였다.<sup>16</sup> 엉덩이 둘레는 엉덩이의 가장 튀어나온 부위의 둘레를 측정하였고 머리 둘레는 앞으로는 미간을 뒤로는 후두골에서 뒤쪽으로 가장 튀어나온 후두부 용기를 지나도록 하여 측정하였다.<sup>17</sup> 엡워스 졸음 척도(Epworth Sleepiness Scale, ESS), 벡 우울 척도(Beck Depression Inventory) 역시 수면다원검사 직전에 시행한 자료를 확인하였고, 당시 환자가 작성한 수면 설문지 및 의무기록을 토대로 고혈압, 당뇨 및 심장질환 이환 여부도 조사하였다. 수면다원검사에서는 OSA 여부 및 정도를 나타내는 AHI값을 확인하였다.

### 무호흡-저호흡 지수(AHI)

무호흡과 저호흡 정의와 진단은 국제수면질환분류 3판(International Classification of Sleep Disorders-3)의 분류 기준을 따랐다.<sup>18</sup> OSA의 진단은 복부와 흉부의 호흡 노력이 관찰되는데도 최소 10초간 구강과 비강에서 무호흡 또는 저호흡이 시간 당 5회 이상 발생하는 상태로 정의하였다. OSA의 중증도 구분은 시간당 발생하는 무호흡 또는 저호흡 횟수에 따른 무호흡-저호흡 지수(AHI)를 기준으로 하여 AHI < 5일 때 정상, 5 ≤ AHI < 15 경도, 15 ≤ AHI < 30 중등도, 그리고 AHI ≥ 30을 중증도로 분류하였다.<sup>19</sup> 무호흡과 저호흡은 미국수면학회(American Association Sleep Medicine)의 기준으로, 무호흡은 공기 흐름이 90% 이상 감소된 상태가 지속적으로 10초 이상 관찰될 때로 정의하였고, 저호흡은 공기 흐름이 50~90%로 감소된 상태가 지속적으로 10초 이상 관찰되며 혈중 산소포화도가 4% 이상 감소한 경우로 정의하였다. 무호흡-저호흡 지수의 경우 시간당 무호흡과 저호흡이 관찰되는 횟수를 계산하여 산정하였다.

### 통계 분석 및 예측 모형

실제 수면다원검사로 측정한 AHI값으로 진단한 경우와 머신 러닝 알고리즘으로 예측한 결과치를 비교하였다. OSA 예측 모델 개발을 위하여 1,368건의 검사를 1,000개의 훈련 자료(training set)와 368의 검증 자료(test set)으로 나누었다. 훈련 자료와 검증 자료의 양 균의 데이터에 유의한 차이가 있는지 검증을 시행하였다. 연속형 변수들은 Shapiro-Wilks 검정을 사용하여 정규분포 여부를 분석하고, 정규분포를 따르는 경우는 평균±표준편차로 표시하고 독립 표본 t-검정을 시행하였으며, 정규분포를 따르지 않는 경우는 중위수(사분위수 범위)로 표시하고 Mann-Whitney 검정을 시행하였다. 범주형 변수는 %(명)으로 표시하였으며, Fisher의 정확 검정을 시행하였다.

다양한 설명 변수를 활용하여 최적의 OSA 예측 모형을 도출하기에 가장 적합할 것으로 기대되는 5가지 머신 러닝 알고리즘, 즉 support vector machine(SVM), lasso logit(LL), naïve bayes(NB), discriminant analysis(DA), K-nearest neighbor(KNN)를 적용하였다. 추정을 위한 모든 데이터는 실측 값에서 평균을 차감한 후, 표준편차로 나누어 정규화하였다. 그리고 모형은 추정의 정확도 향상을 위해 훈련 자료를 겹치지 않게 임의로 10등분하여, 10-겹(fold) 교차 검증(cross validation)을 수행하였으며, MATLAB의 machine learning tool box([https://kr.mathworks.com/products/matlab.html?s\\_tid=hp\\_ff\\_p\\_matlab](https://kr.mathworks.com/products/matlab.html?s_tid=hp_ff_p_matlab))를 활용하여 추정하였다.

### SVM

SVM은 지도 학습(supervised learning) 중 하나로 다양한 환경에서 좋은 성능을 보여주며, 특별한 조정 없이 활용할 수 있는 가장 대중적인 머신 러닝 알고리즘 중 하나이다.<sup>12,20</sup> SVM은 서로 다른 두 집단 사이의 마진(margin)을 최대화하는 최적의 선형 분리선인 초평면을 찾는 알고리즘으로 분류와 회귀 작업 시 유용하다. 고도로 복잡한 자료를 분류하기 위해 자료를 기능별로 다차원으로 매핑(mapping)한다. p-차원의 공간에 n개의 점으로 구성된  $D = \{(x_i, y_i), i=1, 2, \dots, n\}$ 가 존재한다고 하자. 이때 특성을 나타내는  $y_i$ 는 -1 또는 1의 값을 갖는다고 할 때, 초평면  $h(x)$ 는 p-차원에서 아래의 수식(1)과 같이 판별 함수를 제공하며,  $y_i$ 의 특성에 따라 2개로 분류한다.

$$h(x) = W^T x + b = w_1 x_1 + w_2 x_2 + \dots + w_p x_p. \quad (1)$$

수식(1)에서  $w$ 는 p-차원의 가중치 벡터이고  $b$ 는 편향(bias) 상수이다. 가중치 벡터  $W$ 는 초평면에 직교하므로 초평면  $h(x)=0$ 이 주어질 때, 각  $x_i$ 와 초평면 간의 거리  $\delta_i$ 는  $\delta_i = \frac{y_i h(x)}{\|W\|}$

로 나타낼 수 있고, 선형분류기의 마진은 모든 n개의 점에서 초평면까지 최소 거리로 정의된다. 이때 최소 거리를 가지는 모든  $x_i^*$ 를 서포트 벡터(support vectors)라고 한다. SVM의 기본 아이디어는 최대 마진을 가지는 초평면을 선택하는 것으로 아래의 수식(2)와 같이 선형 제약하의 최적화 문제를 푸는 것과 동일하다.

$$\min \frac{\|W\|^2}{2}, \text{ s.t. } y_i h(x_i) \geq 1, \forall x_i \in P. \quad (2)$$

위의 최적화 문제는 라그랑지안 승수법(Lagrange multiplier method)을 활용하여 해결한다.

### LL

LL은 다양한 설명 변수 중 예측을 위한 최적 변수 조합을 도출하는 방법론이다.<sup>12,20</sup> 로지스틱스 회귀 모형에서 종속변수(Y)가 0과 1, 두 가지 값만 취할 때, p개의 독립 변수(X)에 대하여 종속 변수가 1일 확률을 아래의 수식(3)과 같이 나타낼 수 있다.

$$\pi_i (y_i = 1 | x_1, \dots, x_p) = \frac{\exp(\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p)}{1 + \exp(\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p)}. \quad (3)$$

이때  $\beta$ 는 로그우도함수  $\log L(\beta) = -\sum_{i=1}^N [y_i \log(\pi_i) + (1-y_i) \log(1-\pi_i)]$ 를 극소화시켜주는 값이다. LL은 설명 변수의 수가 많은 경우 과적합 문제가 발생하기 때문에 일정한 벌칙(penalty)을 부과하여 최적의 모형을 도출한다. 위 수식(3)의 로그우도함수에 아래의 수식(4)와 같은 벌칙을 부과하여 LL을 추정할 수 있다.

$$\beta^{\text{LASSO}} = \min_{\beta} \left( \log L(\beta) + \lambda \sum_{i=1}^p |\beta_i| \right). \quad (4)$$

이때 수식  $\lambda \sum_{i=1}^p |\beta_i|$ 은 벌칙 항으로  $\lambda$ 의 값이 클수록 벌칙이 강하게 부과되고, 이에 따라 선택되는 설명 변수의 수가 줄어든다. 만약  $\lambda=0$ 을 만족하는 경우에 로그우도함수를 극소화시켜주는  $\beta$ 와 같은 추정치를 얻게 된다. 본 연구에서는 OSA 환자 판별을 위해  $\pi_i$ 가 50% 이상인 경우를 OSA 환자로 정의하였다.

또한, 각 변수들의 OSA와 연관된 중요도 평가를 위해  $AHI \geq 5$ ,  $AHI \geq 15$ ,  $AHI \geq 30$  각각의 경우에서 변수별  $\beta$ 값을 구하였다.  $\beta$ 값이 양수이면 절댓값이 클수록 해당 변수의 값이 커질 때 OSA로 진단받을 가능성이 높아짐을 의미하고,  $\beta$ 값이 음수인 경우는 절댓값이 클수록 해당 변수의 값이 커지면 해당 변수의 값이 커질 때 OSA로 진단받을 확률이 낮아짐을 의미한다.

## NB

NB는 최근에 많이 사용되는 방법으로, 모든 데이터의 특징들이 독립적이라는 가정하에 분류를 위해 베이즈 정리(Bayes' theorem)를 활용한다. 분류를 위한 p-차원의 데이터의 특성을 X, 각 데이터의 분류값을 Y라고 할 때, 새로운 데이터 X가 분류 Y에 속할 확률은 베이즈 정리를 활용하여 아래의 수식(5)와 같이 표현할 수 있다.

$$p(Y|X) = \frac{p(X|Y)p(Y)}{p(X)}. \quad (5)$$

위의 수식(5)는 독립성의 가정을 통해 아래의 수식(6)과 같이 특성의 확률에 대한 곱으로 표현할 수 있다.

$$p(Y=k|x_1, x_2, \dots, x_p) \propto \prod_{i=1}^p p(Y=k)p(x_i|Y=k). \quad (6)$$

수식(6)은 각 변수들이 k값을 갖는 속성에 속할 확률을 나타내고, 이 확률이 가장 큰 값을 갖도록 하는 k를 찾는 것이 NB 알고리즘이다.

## DA

DA는 이미 알려진 두 개 이상의 집단에 속한 관찰값으로부터 각 집단의 차이를 설명하여 줄 수 있는 독립 변수의 선형 결합을 찾아, 알려지지 않은 새로운 개체를 분류하는 기법이다. 두 집단의 확률밀도함수  $f_1(x)$ 가 평균은  $\mu_1$ 이고, 분산-공분산행렬이  $\Sigma$ 로 동일하다고 할 때, 두 확률밀도함수의 비에 로그를 취하면 아래와 같은 수식(7)로 나타낼 수 있다.

$$\begin{aligned} L(x) &= \log \left[ \frac{f_1(x)}{f_2(x)} \right] = \log f_1(x) - \log f_2(x) \\ &= (\mu_1 - \mu_2)' \Sigma^{-1} x - \frac{1}{2} (\mu_1 - \mu_2)' \Sigma^{-1} (\mu_1 - \mu_2). \end{aligned} \quad (7)$$

모집단의  $\mu_1$ ,  $\mu_2$ ,  $\Sigma$ 를 표본으로부터 추정하여  $\bar{x}_1$ ,  $\bar{x}_2$ , S로 나타내면 판별함수를 아래의 수식(8)과 같이 나타낼 수 있다.

$$\hat{L}(x) = (\bar{x}_1 - \bar{x}_2)' S^{-1} x - \frac{1}{2} (\bar{x}_1 - \bar{x}_1)' S^{-1} (\bar{x}_1 - \bar{x}_2). \quad (8)$$

각 집단의 사전 확률을  $\pi_i$ 라고 하면,  $\hat{L}(x_0) \geq \frac{\pi_2}{\pi_1}$ 인 관측치  $x_0$ 는 집단1로 분류하고, 그 이외의 경우에는 집단2에 분류하게 된다.

## KNN

KNN은 가장 널리 사용되는 알고리즘 중 하나로 다양한 방면에서 이용된다. 다양한 유형의 데이터를 처리하기 위한

통계적 패턴인식 방법이다. 새로운 데이터가 주어졌을 때, 기존 훈련 데이터에서 가장 가까운 K개의 데이터를 정보로 새로운 데이터의 속성을 분류하여 학습하는 예측 알고리즘이다. 이때 데이터 점들의 거리는 유클리디안 거리(Euclidean distance)를 활용하는데, 단순히 두 점의 직선거리로 점 X와 Y좌표가 각각  $(x_1, x_2, \dots, x_p)$ ,  $(y_1, y_2, \dots, y_p)$ 로 주어졌을 때, 아래의 수식(9)로 나타낼 수 있다.

$$\text{Euclidean distance}(X, Y) = \sqrt{\sum_{i=1}^p (x_i - y_i)^2}. \quad (9)$$

이때 K값이 매우 작은 경우 과적합(overfitting)의 문제가 발생하고 매우 큰 경우에 오류의 발생가능성이 크기 때문에,<sup>20</sup> 본 연구에서는 K=3으로 설정하였다.

각 알고리즘을 AHI  $\geq 5$ , AHI  $\geq 15$ , AHI  $\geq 30$  세 가지 상황에 적용하였을 때 각각의 정확도(accuracy), 민감도(sensitivity), 특이도(specificity), 정밀도(precision), area under the receiver operating characteristic curve(AUC)를 확인하였다. 이를 위해 true positive(TP), false negative(FN), true negative(TN)와 false positive(FP)를 아래와 같이 정의하고 이를 토대로 지표들을 계산하였다.

TP: AHI값이 정한 범주에 속한 것을 알고리즘에서 예측한 경우.

FN: AHI값이 정한 범주에 속한 것을 알고리즘에서 예측하지 못했을 경우.

TN: AHI값이 정한 범주에 속하지 못한 것을 알고리즘에서 예측한 경우.

FP: AHI값이 정한 범주에 속하지 못한 것을 알고리즘에서 예측하지 못했을 경우.

정확도(accuracy) =  $(TP+TN)/(TP+TN+FP+FN) \times 100$ .

민감도(sensitivity) =  $(TP)/(TP+FN) \times 100$ .

특이도(specificity) =  $(TN)/(TN+FP) \times 100$ .

정밀도(precision) =  $(TP)/(TP+FP) \times 100$ .

## 결 과

### 인구통계학적 특성 및 수면무호흡 정도

전체 평균 연령은  $48.0 \pm 13.1$ 세였고, 성비는 남성이 78.5%, 여성은 21.5%로 남성이 많았다. BMI는  $26.2 \pm 4.3$  kg/m<sup>2</sup>, AHI의 평균은  $28.7 \pm 28.5$ 였다. 고혈압, 당뇨, 심장질환의 기왕력이 있는 경우가 각각 79.9%, 5.6%, 3.7%로 조사되었다(Table 1). 훈련 자료 1,000명 중 AHI  $\geq 5$ 인 경우는 757명(75.7%)이었고, 검증 자료 368명 중 AHI  $\geq 5$ 인 경우는 270명

(73.4%)이 포함되었다. 훈련 자료와 검증 자료를 무작위로 나누었고 각 변수에서 두 자료 간의 유의한 차이가 없는 것을 확인하였다(Table 2).

1,368건의 수면다원검사 중 AHI값에 따라 수면무호흡이 없는, 정상인 결과를 보인 경우가 24.9%로 나타났다. 20.3%

**Table 1.** Demographic characteristics of the included subjects

Feature	Mean	Standard error	Range
Age, year	48.0	13.1	19–82
Height, cm	168.3	8.3	132.0–197.4
Weight, kg	74.7	15.3	37.9–160.0
Body mass index, kg/m <sup>2</sup>	26.2	4.3	15.9–57.4
Head circumference, cm	56.1	2.1	40.0–67.0
Neck circumference (sit), cm	37.8	3.4	28.0–58.0
Waist circumference, cm	91.1	10.5	60.0–160.0
Hip circumference, cm	97.7	9.6	75.0–165.0
Epworth Sleepiness Scale	8.2	4.7	0–24
Beck Depression Inventory	11.6	8.3	0–54
Apnea–Hypopnea Index	28.7	28.5	0–200.8
Sex, % (M:F)	78.5:21.5		
Hypertension, % (Y:N)	79.9:20.1		
Diabetes, % (Y:N)	5.6:94.4		
Heart disease, % (Y:N)	3.7:96.3		

M: male, F: female, Y: yes, N: no

**Table 2.** Comparison between training set and test set

Feature	Training set (n=1,000)	Test set (n=368)	<i>p</i>
Age, year	29.13±28.42	27.30±28.68	0.2246
Sex (male), % (n)	79.3 (793)	76.4 (281)	0.2978
Height, cm	168.2 (132.0, 197.4)	168.7 (143.0, 189.6)	0.3624
Weight, kg	74.6 (37.9, 160.0)	75.0 (44.0, 152.7)	0.6769
Body mass index, kg/m <sup>2</sup>	26.2 (15.9, 57.4)	26.2 (17.9, 50.4)	0.9786
Head circumference, cm	56.2 (40.0, 63.0)	56.1 (40.0, 67.0)	0.5391
Neck circumference (sit), cm	37.8 (28.0, 58.0)	37.9 (30.0, 50.0)	0.7290
Waist circumference, cm	91.2 (60.0, 160.0)	91.0 (70.0, 90.0)	0.7218
Hip circumference, cm	97.6 (75.0, 165.0)	97.8 (75.0, 145.0)	0.7615
Epworth Sleepiness Scale	8.1 (0, 24)	8.2 (0, 23)	0.9244
Beck Depression Inventory	11.6 (0, 54)	11.3 (0, 52)	0.5331
AHI	29.1 (0, 200.8)	27.3 (0, 141.2)	0.2932
Hypertension, % (n)	19.7 (197)	21.2 (78)	0.5434
Diabetes, % (n)	5.2 (52)	6.8 (25)	0.2895
Heart disease, % (n)	3.3 (33)	4.9 (18)	0.1969
AHI≥5, % (n)	75.7 (757)	73.37 (270)	0.3773
AHI≥15, % (n)	55.7 (557)	52.17 (192)	0.2456
AHI≥30, % (n)	39.3 (393)	34.51 (127)	0.1058

Data are presented as median (interquartile range) or mean±standard deviation unless otherwise specified. AHI: Apnea–Hypopnea Index

에서 경한 OSA로 나타내고 있었고 중등도 및 중증 OSA는 각각 16.7%와 38.0%로, 중증 OSA가 가장 높은 비율로 확인되었다(Table 3).

### 예측 모델 성능 확인

5가지 알고리즘 전체에서 AHI≥5 기준 시 머신 러닝의 정확도는 67.5~87.0%였고 정밀도는 83.0~91.9%였다. 또한 민감도와 특이도는 각각 66.0~91.0%, 41.1~74.4%로 나타났다(Table 4). AHI≥15 기준에서는 정확도 66.7~79.6%, 정밀도는 67.0~79.0%였고(민감도 70.2~83.3%, 특이도 58.9~75.65%), AHI≥30의 진단은 66.9~79.6의 정확도와 50.3~68.7%의 정밀도(민감도 43.7~71.4%, 특이도 69.1~88.1%)를 보여주어 정확도는 AHI 기준이 낮았을 때와 비슷했으나 AHI 기준을 올릴수록 정밀도가 떨어지는 양상이었다. 개별 모델을 비교하였을 때, SVM 모델에서는 AHI≥5, AHI≥15, AHI≥30을 기준치로 했을 때 각각의 정확도가 78.6%, 67.7%, 73.0%였으며 정밀도는 각각 83.0%, 67.0%, 64.0%였다. 민감도는 89.6%, 75.8%, 43.7%로 나타났고, 특이도는 41.1%, 58.9%, 87.7%로 나타났다. LL 모델은 각각의 AHI 기준에서 79.6%, 69.1%, 74.1%의 정확도와 84.4%, 69.2%, 65.9%의 정밀도를 보여주었다. 민감도는 89.9%, 73.7%, 46.0%로 나타났다. 특이도는 46.7%, 63.9%, 88.1%로 나타났다. NB의 결과값은 타 알고리즘보다 낮아서 정확도는 67.5%, 66.7%, 66.9%로 AHI 기준과

상관없이 비슷하였고, 정밀도는 88.4%, 67.5%, 50.3%로 AHI 기준을 올렸을 때 급격히 감소하였다. 민감도는 66.0%, 70.2%, 62.7%로 확인되었고, 특이도는 72.2%, 62.8%, 69.9%로 확인되었다. 다른 알고리즘은 AHI 기준이 올라갈수록 특이도는 떨어지고 민감도가 올라가는 양상을 보였으나, NB의 경우는 이와 같은 양상은 관찰되지 않았다. DA 알고리즘으로는 79.6%, 69.1%, 73.8%의 정확도와 84.8%, 69.0%, 65.5%의 정밀도를 확인하였다. 민감도는 89.2%, 74.2%, 45.2%로 나타났고, 특이도는 48.9%, 63.3%, 88.1%로 나타났다. KNN을 활용한 예측 모형이 가장 뛰어난 진단 예측치를 보여주었는데, AHI값에 따른 세 가지 상황 모두에서 좋은 결과를 보여주었다. AHI≥5를 기준으로 했을 때 진단의 정확도는 87.0%, 정밀도 91.9%였으며, AHI≥15 기준에서는 정확도 79.6%, 정밀도 79.0%, AHI≥30의 중증 OSA 진단은 정확도와 정밀도가

각각 79.6%, 68.7%였다. 민감도는 91.0%, 83.3%, 71.4%, 특이도는 74.4%, 75.6%, 83.7%로 나타났고, 특이도와 민감도에서도 다른 알고리즘에 비해 높은 수준으로 나타났으며 AHI 기준이 올라가더라도 다른 알고리즘에 비해 특이도가 크게 떨어지지 않는 것을 확인할 수 있었다.

각 AHI에 따른 5가지 알고리즘의 AUC는 AHI≥5인 경우 76.6~81.9%, AHI≥15인 경우 73.3~79.6%, AHI≥30인 경우 70.4~73.8%로 확인되어, 본 연구에서 제시한 머신 러닝 알고리즘은 OSA 진단에 있어서 중증도의 변별력을 가지고 있는 것을 확인할 수 있었다. 특히 KNN은 AHI≥5인 경우 다른 알고리즘에 비해서 receiver operating characteristic curve의 아래쪽 면적이 가장 넓은 것으로 확인할 수 있었고(Fig. 1), AHI 기준에 따라 각각 81.9%, 79.6%, 73.8%로 나타나 다른 알고리즘에 비해 높은 변별력을 보여주었다.

**Table 3.** Severity classes distribution (n=1,368)

Severity class	AHI	Subjects	Ratio (%)
Normal	AHI<5	341	24.9
Mild	5≤AHI<15	278	20.3
Moderate	15≤AHI<30	229	16.7
Severe	AHI>30	520	38.0

AHI: apnea-hypopnea index

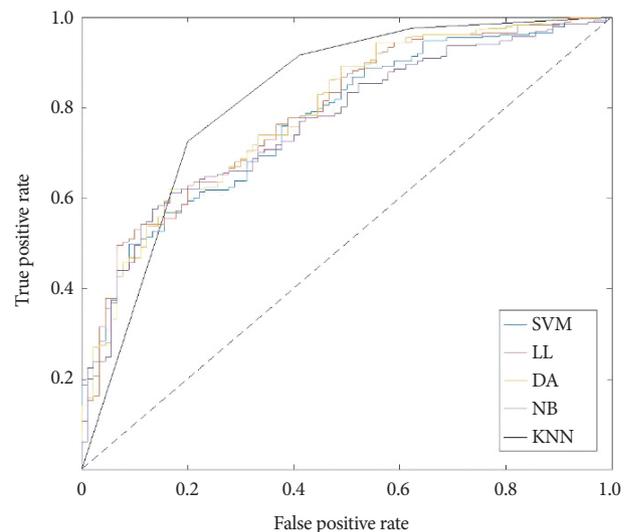
**Table 4.** The performance of the diagnostic models built by different machine learning algorithms

AHI cut-off value	Machine learning methods				
	SVM	LL	NB	DA	KNN
<b>AHI≥5</b>					
Accuracy	78.0	79.6	67.5	79.6	87.0
Sensitivity	89.6	89.9	66.0	89.2	91.0
Specificity	41.1	46.7	72.2	48.9	74.4
Precision	83.0	84.4	88.4	84.8	91.9
AUC	77.2	78.7	76.6	78.9	81.9
<b>AHI≥15</b>					
Accuracy	67.7	69.1	66.7	69.1	79.6
Sensitivity	75.8	73.7	70.2	74.2	83.3
Specificity	58.9	63.9	62.8	63.3	75.6
Precision	67.0	69.2	67.5	69.0	79.0
AUC	77.6	77.2	73.3	76.6	79.6
<b>AHI≥30</b>					
Accuracy	73.0	74.1	66.9	73.8	79.6
Sensitivity	43.7	46.0	62.7	45.2	71.4
Specificity	87.7	88.1	69.1	88.1	83.7
Precision	64.0	65.9	50.3	65.5	68.7
AUC	73.5	73.1	70.4	73.1	73.8

AHI: Apnea Hypopnea Index, AUC: area under the receiver operating characteristic curve, SVM: support vector machine, LL: lasso logit, NB: naïve bayes, DA: discriminant analysis, KNN: K-nearest neighbor

**LL의 변수별 중요도**

LL의 β값으로 각 변수들의 OSA 진단에 대한 중요도를 확인하였을 때, AHI≥5, AHI≥15, AHI≥30 세 경우 모두 나이가 많고 주간 졸림이 많으며 BMI가 높거나 목 둘레, 배 둘레가 길수록 OSA로 진단받을 수 있는 확률이 높아지는 것을 확인하였다(Table 5). AHI≥5, AHI≥30의 경우 당뇨병이 있는 경우 OSA를 진단받을 확률이 높은 것으로 나타났다. 반면, 머리 둘레, 키가 클수록 OSA로 진단받을 확률이 낮으며, 여성일수록 OSA로 진단받을 확률이 낮다는 것을 확인하였다.



**Figure 1.** Area under the receiver operating characteristic curve of prediction models with and AHI cut-off ≥5. ROC: receiver operating characteristic, AHI: Apnea-Hypopnea Index, SVM: support vector machine, LL: lasso logit, DA: discriminant analysis, NB: naïve bayes, KNN: K-nearest neighbor.

**Table 5.** The feature importance in lasso logit

Dependent variables	AHI $\geq$ 5	AHI $\geq$ 15	AHI $\geq$ 30
Age	0.0211	0.0232	0.0097
Body mass index	0.1316	0.1536	0.0667
Head circumference	-0.0278	-0.0523	-
Hip circumference	-	-	-
Height	-0.0120	-0.0177	-0.0037
Neck circumference	0.0343	0.1329	0.1206
Waist circumference	0.0248	0.0079	0.0261
Weight	-	-	-
Epworth Sleepiness Scale	0.0478	0.0599	0.0708
Beck Depression Inventory	-0.0268	-0.0276	-0.0182
Diabetes	0.2204	-	0.2410
Heart disease	-	0.2395	0.0439
Hypertension	-	-0.6368	-0.5245
Sex (male)	-1.3576	-1.0639	-0.5836

If the estimated value is positive, the larger the value, the higher the probability of being diagnosed as OSA when the value of the variable increases. If the estimated value is negative, the larger the value, the lower the probability of being diagnosed as OSA when the value of the variable increases. AHI: Apnea-Hypopnea Index, OSA: obstructive sleep apnea

## 고찰

본 연구에서는 신체 계측치와 간단한 정보들을 이용하여 OSA를 예측하기 위해 머신 러닝 기법을 적용하였다. 머신 러닝 방법론으로 SVM, LL, NB, DA, KNN 5가지 알고리즘을 적용해 보았고 이 중 KNN을 이용한 예측 모델이 가장 좋은 결과를 보여주었다.

OSA를 진단하기 위해서는 수면다원검사가 여전히 유일한 방법이다. 최근 국내에서는 의료보험이 적용되면서 비용 부담이 줄어들긴 하였으나 여전히 고가의 검사이며, 최소 한 명의 검사실 인원이 근무해야 하고, 하룻밤 검사실에서 잠을 자야 한다는 점에서 시공간적인 제한점도 있다. 최근 OSA의 건강에 대한 위험성이 알려지면서 이전에 비해 관심이 증가하고는 있으나 OSA를 방치하고 치료를 하지 않았을 경우 심각한 질병으로 이환될 수 있다는 인식이 부족하며, 여전히 수면클리닉을 방문하는 환자는 전체 환자 비율에 비해 현저하게 낮은 실정이다.<sup>21</sup>

이러한 한계를 극복하고자 수면다원검사를 거치지 않고 사회인구학적 정보와 여러 증상 유무, 그리고 간단한 장비 측정 등으로만 OSA를 예측하기 위한 여러 가지 노력이 있었다.<sup>22-25</sup> 예를 들어, Aaronson 등은 뇌졸중 환자를 대상으로 성별, 나이, BMI 등 환자의 정보와 코골이, 무호흡, 주간 졸림증 등 OSA와 관련된 증상 유무를 이용하여 후진 다중 로

지스틱 회귀분석(backward multivariate logistic regression)을 적용하여 OSA를 선별할 수 있는 예측 모델을 제시하였고, 이 모델은 76%의 정확도를 보여주었다.<sup>26</sup>

최근 활발하게 사용되는 머신 러닝 기법을 이용하면 이러한 예측의 정확도와 정밀도를 향상시킬 수 있다.<sup>27-31</sup> 머신 러닝 기법은 변수들 간의 다차원적인 상호작용을 동시에 고려하여 결론을 도출하는 알고리즘을 제시하기 때문에, 다양한 변수를 하나로 요약할 필요가 없으며 각 변수를 이용한 단일 검정의 한계에도 묶이지 않는 장점을 가지고 있다.<sup>32</sup> OSA의 경우 질환의 중증도와 BMI, 목 둘레, 허리 둘레, 목-허리 둘레 비, 키-허리 둘레 비 및 허리-엉덩이 둘레 비와 같은 신체 계측치가 연관성이 있는 것으로 알려져 있다.<sup>33-36</sup> 이와 같은 인체 계측은 환자 및 정상인에서도 쉽게 측정할 수 있어 OSA 예측 모델을 만드는 변수로 사용하기 유리하다.

기존의 머신 러닝 기법을 적용한 연구들은 대부분 복잡한 수면다원검사 대신 심전도나 광혈류측정기(photoplethysmography)를 이용하여 만든 알고리즘으로 OSA를 진단하는 방법을 제시하는 것으로써,<sup>27,28</sup> 비교적 간단한 검사 장비들을 이용하여 진단할 수 있는 장점이 있다. 하지만 이들 검사 또한 수면시간 동안 직접 착용하여 측정해야만 하는 한계점이 있다. 얼굴 치수, 구강 구조, 아래턱의 움직임 등을 측정하여 OSA를 예측하려는 시도도 있었으나<sup>29,31</sup> 이 또한 측정치를 얻기 위해 안면 촬영(facial scan)을 하거나 아래턱 움직임을 측정할 수 있는 센서를 적용해야 한다는 점에서 번거로움이 있다. 최근 Mencar 등은 나이, 성별, BMI와 함께 OSA 증상, OSA와 관련된 질병의 과거력 및 동맥혈가스분석 및 폐기능 검사(spirometry) 결과치를 이용하여 OSA를 예측하는 모델을 제시하였다.<sup>30</sup> 이 연구에서는 SVM과 random forest 방법을 사용하였고 SVM을 이용한 모델의 정확도는 44.7%, 정밀도는 39.9%였으며, random forest의 경우는 정확도와 정밀도가 모두 44.1%로 나타나 다소 실망스러운 결과를 보여주었다.

과거에도 본 연구처럼 특별한 검사 장비 없이 머신 러닝으로 OSA를 예측하고자 하는 시도가 있었다. Liu 등은 나이, 성별과 함께 머리, 목, 허리, 엉덩이 둘레와 BMI를 측정치, 그리고 베를린설문지, ESS 점수를 이용하여 SVM 알고리즘을 적용한 OSA 예측 모델을 제시하였다.<sup>37</sup> 이 연구에서 여성은 AHI $\geq$ 15를 기준으로 할 때 정확도가 70.1%, AHI $\geq$ 30은 74.6% 정도였고 남성에서는 각각 65.4%, 71.4%로 나타났다. Neural fuzzy evaluation system을 활용한 연구에서는 성별, 나이, BMI, 흡연력, 목과 허리 둘레와 더불어 잠을 자기 전과 아침의 혈압 수치를 이용하였는데, AHI $\geq$ 15의 중증도 이상의 OSA 진단 정확도가 75.6%, 민감도 77.2%, 특이도 75%로, 동일한 집단의 SVM 모델보다 약간 좋은 결과를 보여주었으

며 본 연구의 KNN 모델 결과값과도 비슷한 수치였다.<sup>38</sup> OSA와 연관된 증상만으로 머신 러닝을 적용한 최근 연구도 있다.<sup>39</sup> 이 연구에서는 231명의 수면무호흡으로 진단받은 환자를 대상으로, 코골이, 야뇨증, 코골이 소리에 잠에서 깨기, 요통 등 OSA 연관 증상과 BMI만을 이용하여  $AHI \geq 15$ 인 중등도 이상의 OSA를 예측하기 위해 6가지 머신 러닝 알고리즘(NB, LL, KNN, SVM, neural network, random forest)을 적용하였다. 결과적으로 SVM 모델이 94.42%의 가장 높은 민감도를 보여줬으나 특이도가 32.91%밖에 되지 않았고, NB 모델이 59.49%로 가장 높은 특이도를 구현하였다. 이 연구에서 KNN 모델은 정확도 67.54%, 민감도 78.95%, 특이도 45.57%로 본 연구보다 낮은 예측치를 보여주었다.

임상적으로 OSA 선별을 위해 사용되는 대표적인 도구로는 베를린 설문, STOP, STOP-Bang 설문이 있으며, 이중 STOP-Bang 설문이 선별 도구로는 가장 정확도가 높은 것으로 알려져 있다. 기존 문헌 보고에서  $AHI \geq 5$ 을 기준으로 했을 때 민감도는 베를린 설문이 69.0~95.1%, STOP 설문 74.8~97.0%, STOP-Bang 설문이 81.1~97.6%를 보였고, 특이도는 각각 17.2~41.0%, 6.4~50.0%, 4.9~57.1%로 알려져 있다.<sup>40</sup> 본 연구의 머신 러닝 알고리즘들이 AHI를 동일한 절단값 기준으로 확인하였을 때, 민감도 66.0~91.0%, 특이도 41.1~74.4%를 보였다는 점에서 아직까지는 머신 러닝 기법 전반적으로는 설문지에 비해 월등하게 높은 정확도를 보여 주지는 못하였고, 비슷한 수준의 민감도와 조금 높은 특이도를 나타냈다. KNN 알고리즘이 91.0%의 특이도와 74.4%의 민감도를 보여 기존 설문방식보다 높은 정확도를 나타냈지만 임상적으로 의미 있는 수준으로 판단하기엔 무리가 있다.

다양한 머신 러닝 테크닉을 비교한 이러한 연구에서 서로 다른 결과가 나오는 것은 데이터셋으로 활용한 자료의 양과 환자의 특성, 적용한 변수의 종류가 달라서 발생한 것으로 사료된다. 어떠한 변수와 머신 러닝 알고리즘을 적용했을 때 가장 높은 OSA 진단 예측이 가능한지는 좀 더 많은 연구가 진행되어야 알 수 있을 것이다.

본 연구에서는 다양한 머신 러닝 기법을 이용해 임상에서 손쉽게 얻을 수 있는 정보와 간단한 설문지, 신체 계측만으로 OSA를 진단할 수 있는 예측 모델을 제시하였으며 특히 KNN 기법을 사용하였을 때 가장 좋은 성능을 확인할 수 있었다. 이 예측 모델을 이용하면 수면무호흡을 호소하지 않거나 수면클리닉을 방문할 예정이 없던 환자들에게도 적용하여 OSA의 존재 가능성을 제시함으로써 수면다원검사를 통해 정확한 진단과 향후 적절한 치료를 받을 수 있게 하는 데 도움을 줄 수 있다. 또한 처음으로 한국인을 대상으로 하여 연구를 진행하였다는 점에서 의미가 있다.

하지만 본 연구에서도 여러 가지 제한점이 존재한다. 첫째, 하룻밤의 수면다원검사서 얻은 AHI를 기준으로 진단 여부를 판단하였기에 실제 무호흡 진단율보다는 수면다원검사와의 진단 일치율의 개념이 더 정확할 것이다. 수면다원검사를 처음 받을 때 잠을 잘 이루지 못하는 first night effect를 배제하기 위해 총 수면시간이 4시간 이하인 경우는 제외하였으나 이것 만으로는 부족할 수 있다. 두 번째, OSA에 큰 영향을 미칠 수 있는 modified Mallampati score 및 편도의 크기의 정도를 나타내는 Friedman 병기분류를 적용하지 못하였다.<sup>41,42</sup> 이들 점수 또한 OSA 여부와 관련되어 있는 것으로 알려져 있으나<sup>36,43,44</sup> 검사자가 시각적으로 판단하여 주관적으로 점수를 매기는 방식으로 검사자에 따라 차이가 날 수 있으며, 본 데이터셋에서는 결측치 또한 많았기 때문에 이번 연구에서는 변수로 사용하지 않았다. 마지막으로, 후향적 연구라는 점과 단일 기관의 자료라는 한계점으로 인해 세부 그룹별로 분석하기에는 최종 자료 수가 적었다. 전향적 다기관 연구를 통해 데이터가 더 많이 축적되면 다양한 분석이 가능할 것으로 기대된다.

향후 머신 러닝 기법을 적용하여 좀 더 높은 정확도와 정밀도를 얻기 위해서는 많은 자료가 필수적이므로, 전향적 연구를 통해 다기관 방대한 자료가 추가되고 다양한 세부 집단에서의 검증이 이루어 진다면 더 정확한 OSA 예측 모델을 제시할 수 있을 것으로 기대된다.

## Acknowledgments

This work was supported by clinical research grant from Pusan National University Hospital in 2020.

## Conflicts of Interest

The authors have no potential conflicts of interest to disclose.

## ORCID iDs

Hyun-Woo Kim	<a href="https://orcid.org/0000-0003-1653-7737">https://orcid.org/0000-0003-1653-7737</a>
Euihwan Park	<a href="https://orcid.org/0000-0002-5248-3863">https://orcid.org/0000-0002-5248-3863</a>
Dae Jin Kim	<a href="https://orcid.org/0000-0002-8296-4000">https://orcid.org/0000-0002-8296-4000</a>
Sue Jean Mun	<a href="https://orcid.org/0000-0003-4486-7178">https://orcid.org/0000-0003-4486-7178</a>
Jiyoung Kim	<a href="https://orcid.org/0000-0001-7592-2921">https://orcid.org/0000-0001-7592-2921</a>
Gha-Hyun Lee	<a href="https://orcid.org/0000-0001-9305-9076">https://orcid.org/0000-0001-9305-9076</a>
Jae Wook Cho	<a href="https://orcid.org/0000-0002-2742-9136">https://orcid.org/0000-0002-2742-9136</a>

## Author Contributions

Conceptualization: Hyun-Woo Kim, Euihwan Park. Data curation: Euihwan Park, Dae Jin Kim, Sue Jean Mun. Formal analysis: Hyun-Woo Kim, Euihwan Park, Jiyoung Kim, Gha-Hyun Lee. Methodology: Hyun-Woo Kim, Jiyoung Kim, Jae Wook Cho. Supervision: Hyun-Woo Kim, Jae Wook Cho. Writing—original draft: Hyun-Woo Kim, Euihwan Park, Jae Wook Cho. Writing—review & editing: Hyun-Woo Kim, Jae Wook Cho.

## REFERENCES

1. Peppard PE, Young T, Barnet JH, Palta M, Hagen EW, Hla KM. Increased prevalence of sleep-disordered breathing in adults. *Am J Epidemiol*. 2004;160:5-13.

- demiol* 2013;177:1006-1014. <https://doi.org/10.1093/aje/kws342>.
2. Young T, Palta M, Dempsey J, Peppard PE, Nieto FJ, Hla KM. Burden of sleep apnea: rationale, design, and major findings of the Wisconsin Sleep Cohort study. *WJM* 2009;108:246-249. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2858234/>.
  3. Malhotra A, White DP. Obstructive sleep apnoea. *Lancet* 2002;360:237-245. [https://doi.org/10.1016/S0140-6736\(02\)09464-3](https://doi.org/10.1016/S0140-6736(02)09464-3).
  4. Sunwoo JS, Hwangbo Y, Kim WJ, Chu MK, Yun CH, Yang KI. Prevalence, sleep characteristics, and comorbidities in a population at high risk for obstructive sleep apnea: a nationwide questionnaire study in South Korea. *PLoS One* 2018;13:e0193549. <https://doi.org/10.1371/journal.pone.0193549>.
  5. Nieto FJ, Young TB, Lind BK, et al. Association of sleep-disordered breathing, sleep apnea, and hypertension in a large community-based study. *JAMA* 2000;283:1829-1836. <https://doi.org/10.1001/jama.283.14.1829>.
  6. Shahar E, Whitney CW, Redline S, et al. Sleep-disordered breathing and cardiovascular disease: cross-sectional results of the Sleep Heart Health Study. *Am J Respir Crit Care Med* 2001;163:19-25. <https://doi.org/10.1164/ajrccm.163.1.2001008>.
  7. Engleman HM, Douglas NJ. Sleep · 4: sleepiness, cognitive function, and quality of life in obstructive sleep apnoea/hypopnoea syndrome. *Thorax* 2004;59:618-622. <https://doi.org/10.1136/thx.2003.015867>.
  8. Costa LE, Uchôa CH, Harmon RR, Bortolotto LA, Lorenzi-Filho G, Drager LF. Potential underdiagnosis of obstructive sleep apnoea in the cardiology outpatient setting. *Heart* 2015;101:1288-1292. <https://doi.org/10.1136/heartjnl-2014-307276>.
  9. Kapur V, Strohl KP, Redline S, Iber C, O'Connor G, Nieto J. Underdiagnosis of sleep apnea syndrome in U.S. communities. *Sleep Breath* 2002;6:49-54. <https://doi.org/10.1007/s11325-002-0049-5>.
  10. Mirrakhimov AE, Sooronbaev T, Mirrakhimov EM. Prevalence of obstructive sleep apnea in Asian adults: a systematic review of the literature. *BMC Pulm Med* 2013;13:10. <https://doi.org/10.1186/1471-2466-13-10>.
  11. Li KK, Kushida C, Powell NB, Riley RW, Guilleminault C. Obstructive sleep apnea syndrome: a comparison between Far-East Asian and white men. *Laryngoscope* 2000;110(10 Pt 1):1689-1693. <https://doi.org/10.1097/00005537-200010000-00022>.
  12. Hastie T, Tibshirani R, Friedman J. *The elements of statistical learning: data mining, inference, and prediction*. 2nd ed. New York: Springer, 2009. <https://doi.org/10.1007/978-0-387-84858-7>.
  13. Rajkomar A, Dean J, Kohane I. Machine learning in medicine. *N Engl J Med* 2019;380:1347-1358. <https://doi.org/10.1056/NEJMr1814259>.
  14. Samant P, Agarwal R. Machine learning techniques for medical diagnosis of diabetes using iris images. *Comput Methods Programs Biomed* 2018;157:121-128. <https://doi.org/10.1016/j.cmpb.2018.01.004>.
  15. Gunčar G, Kukar M, Notar M, et al. An application of machine learning to haematological diagnosis. *Sci Rep* 2018;8:411. <https://doi.org/10.1038/s41598-017-18564-8>.
  16. Yang YC, Song EM, Kim KW, Cho JH, Song MY. A study on relationship between lumbosacral curvature and neck-waist circumference on college students in Seoul. *Korean Med Rehab* 2013; 23:168-176.
  17. Cameron N. *Human growth: its assessment, evaluation and variation [dissertation]*. Loughborough: Loughborough University, 2018.
  18. Sateia MJ. International classification of sleep disorders-third edition. *Chest* 2014;146:1387-1394. <https://doi.org/10.1378/chest.14-0970>.
  19. American Academy of Sleep Medicine Task Force. Sleep-related breathing disorders in adults: recommendations for syndrome definition and measurement techniques in clinical research. The Report of an American Academy of Sleep Medicine Task Force. *Sleep* 1999;22:667-689. <https://doi.org/10.1093/sleep/22.5.667>.
  20. James G, Witten D, Hastie T, Tibshirani R. *An introduction to statistical learning with applications in R*. New York: Springer, 2013. <https://doi.org/10.1007/978-1-4614-7138-7>.
  21. Kapur V, Blough DK, Sandblom RE, et al. The medical cost of undiagnosed sleep apnea. *Sleep* 1999;22:749-755. <https://doi.org/10.1093/sleep/22.6.749>.
  22. Farney RJ, Walker BS, Farney RM, Snow GL, Walker JM. The STOP-Bang equivalent model and prediction of severity of obstructive sleep apnea: relation to polysomnographic measurements of the apnea/hypopnea index. *J Clin Sleep Med* 2011;7:459-465. <https://doi.org/10.5664/JCSM.1306>.
  23. Rowley JA, Aboussouan LS, Badr MS. The use of clinical prediction formulas in the evaluation of obstructive sleep apnea. *Sleep* 2000;23:929-938. <https://doi.org/10.1093/sleep/23.7.929>.
  24. Waxman JA, Graupe D, Carley DW. Automated prediction of apnea and hypopnea, using a LAMSTAR artificial neural network. *Am J Respir Crit Care Med* 2010;181:727-733. <https://doi.org/10.1164/rccm.200907-1146OC>.
  25. Deflandre E, Degey S, Brichant JF, Poirrier R, Bonhomme V. Development and validation of a morphologic obstructive sleep apnea prediction score: the DES-OSA score. *Anesth Analg* 2016;122:363-372. <https://doi.org/10.1213/ANE.0000000000001089>.
  26. Aaronson JA, Nachtegaal J, van Bezeij T, et al. Can a prediction model combining self-reported symptoms, sociodemographic and clinical features serve as a reliable first screening method for sleep apnea syndrome in patients with stroke? *Arch Phys Med Rehabil* 2014;95:747-752. <https://doi.org/10.1016/j.apmr.2013.12.011>.
  27. Khandoker AH, Palaniswami M, Karmakar CK. Support vector machines for automated recognition of obstructive sleep apnea syndrome from ECG recordings. *IEEE Trans Inf Technol Biomed* 2009;13:37-48. <https://doi.org/10.1109/ITB.2008.2004495>.
  28. Uçar MK, Bozkurt MR, Bilgin C, Polat K. Automatic sleep staging in obstructive sleep apnea patients using photoplethysmography, heart rate variability signal and machine learning techniques. *Neural Comput & Applic* 2018;29:1-16. <https://doi.org/10.1007/s00521-016-2365-x>.
  29. Islam SMS, Mahmood H, Al-Jumaily AA, Claxton S. Deep learning of facial depth maps for obstructive sleep apnea prediction. *International Conference on Machine Learning and Data Engineering (iCMLDE)*; 2018 December 3-7; Sydney, Australia. IEEE 2019;154-157. <https://doi.org/10.1109/iCMLDE.2018.00036>.
  30. Mencar C, Gallo C, Mantero M, et al. Application of machine learning to predict obstructive sleep apnea syndrome severity. *Health Informatics J* 2020;26:298-317. <https://doi.org/10.1177/1460458218824725>.
  31. Pépin JL, Letesson C, Le-Dong NN, Dedave A, Denison S, Cuthbert V, et al. Assessment of mandibular movement monitoring with machine learning analysis for the diagnosis of obstructive sleep apnea. *JAMA Netw Open* 2020;3:e1919657. <https://doi.org/10.1001/jamanetworkopen.2019.19657>.
  32. Chekroud AM, Zotti RJ, Shehzad Z, et al. Cross-trial prediction of treatment outcome in depression: a machine learning approach. *Lancet Psychiatry* 2016;3:243-250. [https://doi.org/10.1016/S2215-0366\(15\)00471-X](https://doi.org/10.1016/S2215-0366(15)00471-X).
  33. Öğretmenoğlu O, Süslü AE, Yücel OT, Onerci TM, Sahin A. Body fat composition: a predictive factor for obstructive sleep apnea. *Laryngoscope* 2005;115:1493-1498. <https://doi.org/10.1097/01.mlg.0000172204.82314.c3>.
  34. Wysocki J, Charuta A, Kowalcze K, Ptaszyńska-Sarosiek I. Anthropometric and physiologic assessment in sleep apnoea patients regarding body fat distribution. *Folia Morphol* 2016;75:393-399. <https://doi.org/10.5603/FM.a2015.0127>.
  35. Katz SL, Vaccani JP, Barrowman N, Momoli F, Bradbury CL, Murto K. Does neck-to-waist ratio predict obstructive sleep apnea in children? *J Clin Sleep Med* 2014;10:1303-1308. <https://doi.org/10.5664/jcsm.4284>.
  36. Yang KI. [Sleep-related respiratory disorder]. In: Kim JH, Sunwoo J, Song P, et al. *Sleep disorders: case-based learning*. Seoul: PanMun Education, 2020:154-156.
  37. Liu WT, Wu HT, Juang JN, et al. Prediction of the severity of obstructive sleep apnea by anthropometric features via support vector machine. *PLoS One* 2017;12:e0176991. <https://doi.org/10.1371/journal.pone.0176991>.

38. Wu MF, Huang WC, Juang CF, et al. A new method for self-estimation of the severity of obstructive sleep apnea using easily available measurements and neural fuzzy evaluation system. *IEEE J Biomed Health Inform* 2017;21:1524-1532. <https://doi.org/10.1109/JBHI.2016.2633986>.
39. Keshavarz Z, Rezaee R, Nasiri M, Pournik O. Obstructive sleep apnea: a prediction model using supervised machine learning method. *Stud Health Technol Inform* 2020;272:387-390. <https://doi.org/10.3233/SHTI200576>.
40. Yang H, Park H. Usefulness of the Berlin, STOP, and STOP-Bang Questionnaires in the Diagnosis of Obstructive Sleep Apnea. *J Sleep Med* 2019;16:11-20. <https://doi.org/10.13078/jsm.19021>.
41. Nuckton TJ, Glidden DV, Browner WS, Claman DM. Physical examination: Mallampati score as an independent predictor of obstructive sleep apnea. *Sleep* 2006;29:903-908. <https://doi.org/10.1093/sleep/29.7.903>.
42. Friedman M, Salapatas AM, Bonzelaar LB. Updated Friedman Staging System for obstructive sleep apnea. *Adv Otorhinolaryngol* 2017;80:41-48. <https://doi.org/10.1159/000470859>.
43. Friedman M, Hamilton C, Samuelson CG, Lundgren ME, Pott T. Diagnostic value of the Friedman tongue position and Mallampati classification for obstructive sleep apnea: a meta-analysis. *Otolaryngol Head Neck Surg* 2013;148:540-547. <https://doi.org/10.1177/0194599812473413>.
44. Yu JL, Rosen I. Utility of the modified Mallampati grade and Friedman tongue position in the assessment of obstructive sleep apnea. *J Clin Sleep Med* 2020;16:303-308. <https://doi.org/10.5664/jcsm.8188>.