Predictive Screening for Obstructive Sleep Apnea Using Non-Intrusive Patient Data and Anthropometric Measurements

Diego Estuar
destuar@lion.lmu.edu
Loyola Marymount University
Los Angeles, USA

Andrew Xu
cxu2@lion.lmu.edu
Loyola Marymount University
Los Angeles, USA

Yannick Angouo Lopes
yangouol@lion.lmu.edu
Loyola Marymount University
Los Angeles, USA

Abstract

Study Objectives: Evaluate the performance of Logistic Regression, XGBoost, and Multi-Layer Perceptron (MLP) in predicting obstructive sleep apnea (OSA) using non-intrusive patient data and anthropometric measurements, such as neck and waist circumference ratios to height, alongside demographic and lifestyle factors.

Methods: Utilizing data from the Sleep Heart Health Study comprising 5,804 participants with polysomnography results and anthropometric data, this study compares three machine learning models to determine their accuracy in identifying moderate to severe OSA.

Results: The study identifies Logistic Regression as the most effective model, particularly noted for its superior positive class recall performance, which is crucial for minimizing false negatives in clinical screenings. This finding demonstrates the potential of simple, non-invasive measurements to facilitate early and accessible screening for OSA, particularly in settings with limited access to advanced medical diagnostics.

Conclusions: The study's findings reveal that while the predictive models, especially Logistic Regression, show promising results in early OSA screening, further research is necessary to increase the positive class recall and overall performance of all models. Continued improvements and validation across diverse datasets are critical before these models can be applied on a large scale and integrated effectively into clinical practice to reduce healthcare burdens and improve patient outcomes.

Conference'24, April 2024, Los Angeles, CA, USA

ACM ISBN 978-1-4503-XXXX-X/24/04
https://doi.org/XXXXXXX.XXXXXXX

Keywords: Obstructive sleep apnea, machine learning, Logistic Regression, XGBoost, Multi-Layer Perceptron, anthropometry, predictive modeling, healthcare analytics

ACM Reference Format:
Diego Estuar, Andrew Xu, and Yannick Angouo Lopes. 2024. Predictive Screening for Obstructive Sleep Apnea Using Non-Intrusive Patient Data and Anthropometric Measurements. In . ACM, Los Angeles, CA, USA, 5 pages. https://doi.org/XXXXXXX.XXXXXXX

1 Introduction

Obstructive Sleep Apnea (OSA) is a significant health concern affecting a large population of individuals. In some cases, it can also lead to serious comorbidities. OSA primarily affects individuals during sleep with symptoms that are easy to overlook or misinterpret, which makes it particularly challenging to diagnose early. The urgency to identify those at high risk of severe OSA has encouraged research into effective and accessible screening tools. This study investigates the application of machine learning algorithms to predict OSA in patients using anthropometric measurements and other non-intrusive patient features from the Sleep Heart Health Study cohort.

We aim to advance the predictive capabilities by testing three distinct machine learning models to determine which most accurately identifies patients at risk based on their physical measurements alone. The models will assess the effectiveness of using non-intrusive patient data anthropometric data, an easily obtainable and non-invasive measure, to predict the severity of OSA. This approach could significantly enhance early screening processes, especially in settings lacking advanced medical equipment or specialized sleep study facilities.

Data for this research will be drawn from the Sleep Heart Health Study, a well-established source that includes a wide demographic and has been pivotal in previous OSA research. This study's outcomes could lead to more tailored preventive measures and interventions, reducing the health care burdens associated with OSA and improving patient outcomes across diverse populations.

2 Related Work

In the pursuit of non-invasive and accessible screening tools for Obstructive Sleep Apnea (OSA), machine learning (ML) applications stand out for their potential to transform diagnostic procedures through the implementation of algorithmic predictions. OSA is recognized for its significant health implications, including cardiovascular diseases and decreased quality of life, necessitating the importance of accurate and efficient predictive models for early intervention (Shamsuzzaman et al., 2003). The emergence of ML applications using anthropometric data, measurements such as weight, height and size, provide a potential promising solution.

2.1 Sleep Heart Health Study (SHHS)

The Sleep Heart Health Study (SHHS) is a pivotal longitudinal cohort study that has substantially contributed to understanding the cardiovascular and other systemic impacts of sleep-disordered breathing, including OSA. This study has recruited thousands of participants from various existing cardiovascular cohort studies, capitalizing on a diverse population to explore the relationships between sleep, breathing, and heart health over a prolonged period.

The SHHS provides a rich database of full-night comprehensive sleep studies, alongside extensive cardiovascular, respiratory, and metabolic health measurements. The inclusion of anthropometric data such as neck and waist circumferences, as noted by Vana et al. (2021), has allowed for the exploration of physical predictors in relation to sleep apnea. The breadth of the SHHS dataset has enabled models with a high degree of accuracy due to the large sample size and the comprehensive nature of the data, which also includes demographic information, health behaviors, and various other physiological measurements. This comprehensive approach enhances the models' generalizability and reliability in predicting the severity of OSA across diverse populations, demonstrating the potential of similar non-invasive and cost-effective screening tools.

2.2 Machine Learning Applications

Building on the foundational data provided by the Sleep Heart Health Study (SHHS), the advancement and application of machine learning tools in the field of sleep medicine have made significant progress. Kim et al. (2024) integrated craniofacial photographs with questionnaires, yielding an AUROC of 94.1%. The success of Gradient Boosting Machines, as reported by Shi et al. (2023), further validated the potential applications of machine learning in healthcare, especially in early disease detection. Tsai et al. (2022) provided insights into visceral fat's predictability in OSA, demonstrating the relevance of body composition in ML models. Similarly, neural network analyses of ECG data have indicated high diagnostic accuracy, suggesting the viability of ECG-based deep learning models for non-invasive OSA diagnosis (Xingyu Li et al., 2023; Lin et al., 2023). These computational approaches have been explored across different populations and conditions. For example, in COPD patients, ML models have revealed a high prevalence of OSA, suggesting a need for sleep assessments within pulmonary rehabilitation programs (Soler et al., 2015). Additionally, research by Friedman et al. (2009) identified reliable physical predictors for OSA, which can be used to formulate a predictive score.

Recent developments include the design of a nomogram for predicting OSA in patients with chronic coronary syndrome, showcasing discriminative ability (Xu et al., 2022). Such tailored models reinforce the adaptability of ML to various patient demographics, as echoed by studies focusing on gender differences (Appleton et al., 2018) and the impact on quality of life (Moyer et al., 2001). Despite these advancements, challenges remain, particularly in the interpretation of model outputs and handling of imbalanced datasets (Kheirandish-Gozal et al., 2016). However, the continued exploration of ML, with a high index of clinical suspicion, may eventually lead to reduced healthcare burdens associated with OSA and improved patient outcomes (Siriyotha et al., 2021).

2.3 Research Gap

Given the substantial work already conducted using machine learning to enhance diagnostic accuracy for Obstructive Sleep Apnea (OSA), the research gap primarily centers around the application and refinement of specific machine learning algorithms. Prior studies have demonstrated the efficacy of various models, yet there remains significant potential for improving the precision and generalizability of these methods in clinical settings. Specifically, leveraging a unique combination of non-intrusive patient data and advanced machine learning techniques, including Logistic Regression, XGBoost, and Multi-Layer Perceptron (MLP), presents an opportunity to enhance the predictive accuracy of OSA screening tools.

This approach aims to integrate a distinctive set of features derived from anthropometric measurements and other non-invasive data sources to develop a model that not only predicts OSA with a higher positive class recall and AUC-ROC, but also adapts effectively to diverse patient populations. While previous studies have laid a strong foundation, the integration of these specific algorithms with a carefully selected feature set could lead to significant advancements in the non-invasive screening of OSA. This research aims to fill this gap through the fine-tuning of these models and validation of their effectiveness given non-invasive patient data and anthropemtric measurements as features.

3 Data and Methodology

The Sleep Heart Health Study (SHHS) dataset, foundational to our research, comprises data from a longitudinal cohort study launched in 1995 to delve into the associations between sleep-disordered breathing and cardiovascular health. Drawing participants from varied existing cardiovascular cohort studies, the SHHS offers an extensive dataset that features full-night polysomnography, cardiovascular diagnostics, and a wide range of metabolic and respiratory health measurements. Our selection of the SHHS dataset is due to its incorporation of extensive anthropometric and non-intrusive patient data—elements for the development of predictive models for Obstructive Sleep Apnea (OSA). This dataset not only provides a robust sample size but also includes detailed demographic profiles and health behavior records, which are importnat for enhancing the accuracy and generalizability of our predictions. This data infrastructure supports our methodology in applying advanced machine learning techniques to predict OSA, highlighting its potential to serve as a non-invasive, cost-effective tool for early diagnosis and management in diverse populations. Figure 1 presents the database schema of the Sleep Heart Health Study (SHHS) dataset, outlining the structure and relationships between different data entities utilized in our analysis.

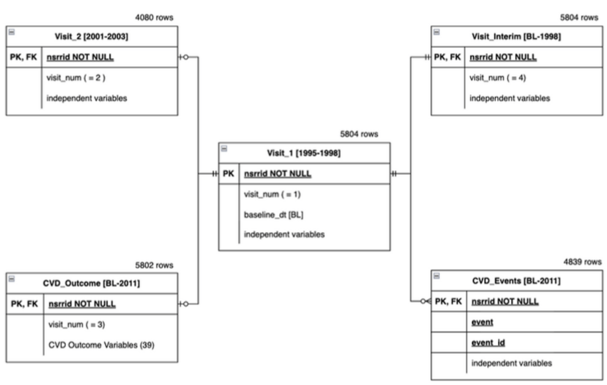

Figure 1. Database schema showing the structure of the Sleep Heart Health Study dataset used in the study.

3.1 Logistic Regression

Given the task of classifying patients with or without OSA using non-invasive data, Logistic Regression was chosen for its proven efficacy in binary classification problems and its ability to provide easily interpretable probabilistic outcomes, which are essential in medical diagnostics. This model's strengths lie in its straightforward implementation and the clarity of the insights it can generate, such as odds ratios for each predictor. However, it assumes linearity between independent variables and the log-odds of the dependent variable, which can limit its effectiveness in capturing complex relationships. The model is also sensitive to outliers and requires a sufficiently large sample size to ensure reliable estimates.

The logistic regression model predicts the probability of Obstructive Sleep Apnea (OSA) as follows:

$$\hat{y} = \sigma(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n) \tag{1}$$

where $\sigma(z) = \frac{1}{1+e^{-z}}$ is the logistic function.

Performance metrics indicate a Test AUC-ROC Score of 0.71, with the model demonstrating notable precision in predicting positive cases of OSA, reflected in a positive class recall of 0.62.

3.2 XGBoost

XGBoost was selected for its capacity to handle diverse and complex datasets, making it optimal for the multi-dimensional nature of non-invasive patient data and anthropometric measures used to predict OSA. Its robustness against overfitting and ability to process nonlinear relationships through a gradient boosting framework allows for the optimization of loss functions and enhanced predictive accuracy. Despite its strengths in performance, XGBoost is often criticized for its "black box" nature, making it challenging to interpret the model's decision-making process. The training time can also be extensive due to its complex ensemble structure.

The XGBoost model is represented by the equation:

$$\hat{y}_i = \sum_{k=1}^{K} f_k(x_i), \quad f_k \in \mathcal{F} \tag{2}$$

where each f_k is a decision tree.

In this study, XGBoost achieved a Test AUC-ROC Score of 0.68, but it showed limitations in detecting positive cases of OSA, with a low positive class recall of 0.15, highlighting an area for model refinement and tuning.

3.3 Multi-Layer Perceptron

The Multi-Layer Perceptron (MLP) was chosen for this research due to its robust capability as a universal function approximator, essential for managing the nonlinear complexities of medical datasets like those involved in obstructive sleep apnea (OSA) prediction. However, the model has limitations and is often viewed as a "black box" due to the difficulty in interpreting its internal workings. There's also a notable risk of overfitting, especially with smaller data sets. Moreover, MLPs require substantial amounts of data to train effectively and necessitate feature scaling to ensure uniform contribution of each input feature.

The MLP model, with three hidden layers each containing 50 nodes and employing the tanh activation function, is designed to optimize the weights using a learning rate of 0.0001. This model structure is represented mathematically as follows:

$$\hat{y} = \text{softmax}(z^{(3)}) \tag{3}$$

$$z^{(3)} = W^{(3)} \cdot a^{(2)} + b^{(3)} \tag{4}$$

$$a^{(2)} = \tanh(z^{(2)}) \tag{5}$$

$$z^{(2)} = W^{(2)} \cdot a^{(1)} + b^{(2)} \tag{6}$$

$$a^{(1)} = \tanh(z^{(1)}) \tag{7}$$

$$z^{(1)} = W^{(1)} \cdot x + b^{(1)} \tag{8}$$

where:

- $W^{(1)}, W^{(2)}, W^{(3)}$ are the weight matrices of the three layers.
- $b^{(1)}, b^{(2)}, b^{(3)}$ are the bias vectors for each layer.
- x is the input vector to the network.
- tanh is the hyperbolic tangent activation function.
- $\text{softmax}(z_i) = \frac{e^{z_i}}{\sum_k e^{z_k}}$ normalizes the output into a probability distribution.
- The learning rate of 0.0001 is used to adjust the weights incrementally during training, ensuring stable convergence.

The performance of the MLP was evaluated, and the metrics presented on the test data show a Test AUC-ROC Score of 0.659. Precision, recall, and F1 scores are broken down by class, where Class 0 displays high precision and recall, indicating an effective prediction of negative cases. In contrast, Class 1 shows significantly lower performance, suggesting that the model struggles with accurately identifying positive cases of OSA. This highlights potential areas for further refinement, particularly in improving the model's sensitivity to positive cases.

4 Results

The study evaluated three predictive models Logistic Regression, XGBoost, and Multi-Layer Perceptron (MLP) for diagnosing Obstructive Sleep Apnea (OSA) using non-intrusive data. The comparative performance of the models in predicting OSA is illustrated in the ROC curves shown in Figure 2, which plot the true positive rate against the false positive rate for each model.

Logistic Regression emerged as the highest performing model with an AUC-ROC of 0.71 and recall of 0.62 for the positive class, indicating its ability to correctly identify cases of OSA. This is particularly significant in clinical settings where high recall is crucial to minimize the risk of missing a diagnosis. In contrast, XGBoost and MLP, despite their competitive AUC scores of 0.68 and 0.66 respectively, show lower recall rates (0.15 and 0.17), which could lead to missed diagnoses. The Receiver Operating Characteristic (ROC) curves further illustrate Logistic Regression's optimal balance of sensitivity and specificity, confirming its effectiveness and reliability in a healthcare context where the cost of a false negative is substantially higher than that of a false positive.

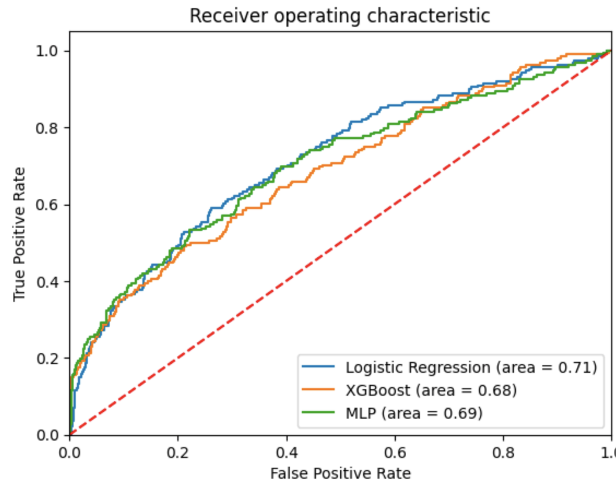

Figure 2. ROC curves for Logistic Regression, XGBoost, and MLP models, displaying the trade-off between true positive rate and false positive rate along with the area under the curve (AUC) for each model.

5 Observations and Conclusions

Our study successfully demonstrated the feasibility of using non-intrusive data to predict OSA. Patient features such as age, gender, neck circumference, waist measurements, average BMI, presence of sleep disorders, instances of stopped breathing, and history of treated snoring are significant predictors within our machine learning models. These models have achieved high accuracy, using data that is not only easily collectible without causing discomfort to patients but also crucial for early diagnosis and intervention. Such early detection is vital in averting severe complications often seen with untreated sleep disorders.

Many opportunities exist for future research to advance the capabilities of predictive models for obstructive sleep apnea (OSA). Subsequent research should aim to broaden the dataset's diversity to enhance the models' applicability and reduce biases. Integrating additional variables such as lifestyle choices, genetic predispositions, and environmental factors will likely increase the diagnostic accuracy. Longitudinal studies are recommended to assess the progression of OSA over time and evaluate the long-term impacts of early interventions. Additionally, conducting longitudinal studies could provide a better understanding of OSA's progression and the long-term efficacy of early interventions, thereby optimizing the predictive algorithms to respond to long-term outcomes. This approach underlines the practicality of using non-intrusive data for predicting OSA and emphasizes the importance of achieving high sensitivity and specificity in diagnostic models.

References

[1] Appleton, S. L., Gill, T. K., Taylor, A. W., McEvoy, R. D., Shi, Z., Hill, C. L., Reynolds, A. C., Adams, R. J. Gender Influence on OSA Symptoms and Quality of Life. *Int. J. Environ. Res. Public Health*, 15, 5 (2018), 930.

[2] Friedman, M., Tanyeri, H., La Rosa, M., Landsberg, R., Vaidyanathan, K., Pieri, S., Caldarelli, D. Clinical Predictors of Obstructive Sleep Apnea. *Laryngoscope*, 109, 12 (1999), 1901-1907.

[3] Kheirandish-Gozal, L., Gozal, D. Pediatric OSAS: Oximetry Can Provide Answers When Questionnaires Cannot. *Sleep Med. Rev.*, 27 (2016), 96–105.

[4] Kim, J., et al. Machine Learning Models for OSA Prediction. *Sleep Research Society*.

[5] Lin, Z., et al. ECG Signal Analysis for OSA Detection via Wavelet Transform and Deep Learning. *Research Square*.

[6] Moyer, C. A., Sonnad, S. S., Garetz, S. L., Helman, J. I., Chervin, R. D. Quality of Life in Obstructive Sleep Apnea: A Systematic Review of the Literature. *Sleep Medicine*, 2, 6 (2001), 477–491.

[7] Shamsuzzaman, A. S. M., Gersh, B. J., Somers, V. K. Obstructive Sleep Apnea: Implications for Cardiac and Vascular Disease. *JAMA*, 290, 14 (2003), 1906–1914.

[8] Shi, X., et al. Machine Learning Models for Severe OSA Prediction. *BMC Medical Informatics and Decision Making*.

[9] Siriyotha, S., Tantrakul, V., Plitphonganphim, S., Rattanasiri, S., Thakkinstian, A. Prediction Models of Obstructive Sleep Apnea in Pregnancy: A Systematic Review and Meta-Analysis of Model Performance. *Diagnostics*, 11, 6 (2021), 1097.

[10] Soler, X., Gaio, E., Powell, F. L., Ramsdell, J. W., Loredo, J. S., Malhotra, A., Ries, A. L. High Prevalence of Obstructive Sleep Apnea in Patients with Moderate to Severe Chronic Obstructive Pulmonary Disease. *AnnalsATS*, 12, 8 (2015), 1219–1225.

[11] Tsai, P. H., et al. Machine Learning and Anthropometric Features for OSA Screening. *Sensors*, 22, 22 (2022), 8630.

[12] Vana, K. D., et al. Anthropometric Measures and OSA in the Sleep Heart Health Study Cohort. *J. Clin. Sleep Med.*

[13] Xingyu Li, et al. Machine Learning Models for OSA Prediction Using ECG Data. *Sci. Rep. (Nature)*.

[14] Xu, Y., Ye, Z., Wang, B., Tang, L., Sun, J., Chen, X., Yang, Y., Wang, J. Novel Insights into the Predictors of Obstructive Sleep Apnea Syndrome in Patients with Chronic Coronary Syndrome: Development of a Predicting Model. *Oxid. Med. Cell. Longev.*