The Importance of Health Informatics in Public Health during a Pandemic
J. Mantas et al. (Eds.)

doi:10.3233/SHTI200576

Obstructive Sleep Apnea: A Prediction Model Using Supervised Machine Learning Method

Zahra KESHAVARZ [a], Rita REZAEE [b,1], Mahdi NASIRI [b] and Omid POURNIK [c]
[a] *Student research committee, School of Management and Medical information Sciences, Shiraz University of Medical Sciences, Shiraz, Iran*
[b] *Health Human Resources Research Center, School of Management & Information Sciences, Shiraz University of Medical Sciences, Shiraz, Iran*
[c] *Department of Community Medicine, School of Medicine, Iran University of Medical Sciences, Tehran, Iran*

Abstract. Obstructive Sleep Apnea (OSA) is the most common breathing-related sleep disorder, leading to increased risk of health problems. In this study, we investigated and evaluated the supervised machine learning methods to predict OSA. We used popular machine learning algorithms to develop the prediction models, using a dataset with non-invasive features containing 231 records. Based on the methodology, the CRISP-DM, the dataset was checked and the blanked data were replaced with average/most frequented items. Then, the popular machine learning algorithms were applied for modeling and the 10-fold cross-validation method was used for performance comparison purposes. The dataset has 231 records, of which 152 (65.8%) were diagnosed with OSA. The majority was male (143, 61.9%). The results showed that the best prediction model with an overall AUC reached the Naïve Bayes and Logistic Regression classifier with 0.768 and 0.761, respectively. The SVM with 93.42% sensitivity and the Naïve Bayes of 59.49% specificity can be suitable for screening high-risk people with OSA. The machine learning methods with easily available features had adequate power of discrimination, and physicians can screen high-risk OSA as a supplementary tool.

Keywords. Obstructive Sleep Apnea, Prediction, Data Mining, Supervised Machine Learning Methods

1. Introduction

Obstructive sleep apnea (OSA) is the most common breathing-related sleep disorder, affecting 9-38% of the world population [1]. The OSA has serious and unpleasant consequences. If this disorder is not diagnosed or treated, there will be serious health problems and significant economic burden for both the individual and the healthcare system. The gold standard to diagnose and determine OSA severity is polysomnography (PSG). The PSG is a medical procedure performed in a specialized sleep laboratory by monitoring various aspects of the body activity during sleep. Indeed,

[1] Corresponding author, Dr. Rita Rezaee, Clinical Education Research Center, Health Human Resources Research Center, School of Management & Information Sciences, Shiraz University of Medical Sciences, Shiraz, Iran; Email: Rezaeerita@yahoo.com.

these can be time consuming and costly [2]. Portable sleep monitoring devices are common replacement for PSG test. Although there is less work required in the lab, the patient should have some certain conditions [3]. Many studies used ECG, EEG, SpO2 or the combination of two or more signals to diagnose OSA. Although these methods have shown high accuracy and good ability to predict OSA, they merely focus on procedures that require a specialist and equipment for recording and analyzing OSA [4]. Recently, many studies applied machine learning methods in medical fields including sleep medicine to classify people into healthy or ill. In a systematic review by Pombo et al., it was found that amongst the classification methods used for predicting OSA, the machine learning (ML) at 85.25% was the most used method and most popular algorithm which has good performance and reliability [5]. Due to the need for an efficient method that is simple, needs less effort and has good sensitivity for diagnosing and screening of high-risk people, in the present study we used a dataset containing self-reported variables which can be obtained effortlessly by utilizing the CRISP-DM (Cross-industry standard process for data mining) instruction, a methodology for the data mining project, in order to present a model for OSA diagnosis.

2. Methods

This study was approved by the local Ethics Committee of Shiraz University of Medical Sciences with the ethics code of IR.SUMS.REC.1396.3885. Written informed consents were obtained before performing the PSG test. To commence this study, we used CRISP-DM methodology which is a standard approach for planning a data mining project in the medical field. The steps are described as below. In the first phase, business understanding, the objective was to develop a model for predicting OSA in order to select the best model for assisting physicians and medical staff to determine and screen high-risk OSA patients. In the next phase, data understanding and preparation, the focus was on data. The data were collected retrospectively from medical records of people that had performed PSG test at Nemazi Hospital affiliated with Shiraz University of Medical Sciences from February 2013 to December 2017. The used dataset has 250 records and 49 features. The AHI was divided into two classes (OSA and healthy) with the threshold of 15[6]. At first, we eliminated 19 samples who were under 18 years of age, diagnosed with insomnia or most of the features were not reported. Records containing missing values were replaced with average/most frequented items. We used rank widget with information gain method as feature selection techniques through Orange (v.2), a machine learning free toolkit. The rank widget considers class-labeled data sets and scores the features according to their correlation with each class. In the phase of modeling to predict the accurate OSA diagnosis, the best-ranked features with information gain method were entered as inputs for the Neural Networks, Naïve Bayes, Logistic Regression, k-nearest neighbor (KNN), support vector machine (SVM) and Random Forest algorithms for training. The training of all algorithms was conducted using the same features. In the phase of evaluation and development, in order to evaluate the models, 10-fold cross-validation method was used. To compare the performance between prediction models accuracy, the area under the ROC curve (AUC), sensitivity and specificity of the models were evaluated.

3. Results

The final dataset had 231 records, of which 152 (65.8%) were diagnosed as OSA based on the threshold 15 of AHI. The majority of people were male (143, 61.9%). The variables including severe snore, nocturia, awakening due to the sound of snoring, witnessed snore, witnessed apnea, back pain, restless sleep, and BMI that have important roles in prediction of OSA based on the rank score were fed as the input of the models. In addition, the models were evaluated with other selected important features; therefore, they had no significance. In the assessed models, the parameters of accuracy, AUC, sensitivity and specificity of each machine learning algorithms are shown in Table 1, to compare the performance of prediction models. The SVM with 93.42% sensitivity had predicted the majority of OSA patients better, and the Naïve Bayes with 59.49% specificity had predicted the healthy people better than the other models. The Naïve Bayes and Logistic Regression classifier had the highest AUC with 0.768 and 0.761, respectively (Table 1). According to the results, the neural network had a better performance in the assessed models.

Table 1. The results of 10-fold cross validation of the mode

Method	CA	Sens	Spec	AUC
Naive Bayes	72.30	78.95	59.49	0.7681
Logistic regression	73.61	86.18	49.37	0.7614
Neural Network	74.91	86.18	53.16	0.7528
KNN	67.54	78.95	45.57	0.6557
SVM	72.77	94.42	32.91	0.7225
Random Forest	74.74	87.50	40.51	0.7586

Note: CA: Classification Accuracy, Sens: sensitivity, Spec: specificity, AUC: area under the ROC curve

4. Discussion

In the present study, to present a simple prediction model, the supervised machine learning methods with a standard methodology was used to predict OSA. In this study by selecting the appropriate method of feature selection, less important features were removed and models had a high accuracy, which were in accordance to the studies of Aronson et al. [7]. The goal was to ensure that high-risk OSA people will not be left without a PSG test; therefore, in order to reach a high sensitive model to avoid false negatives high sensitivity was important. Amongst the trained models, the SVM sensitivity reached a good result in determining high-risk OSA patients. This result was higher than the studies of Wu et al., and Liu et al. [6,8]. Specificity represents the probability of a negative result in people without the disease. According to the results, the Naïve Bayes with the 59.49% specificity had the highest ability to identify people with low risk of OSA, which was in line with the studies of Leite et al. and Aaronsan et al., but it was higher than the study by Wu et al [6, 7, 9]. To the best of our knowledge, there is no other study with developed models, using self-reported symptoms and easily available features with machine learning methods for predicting the presence of OSA in people. Researchers successfully applied machine learning to predict OSA patients, but the majority of these prediction models need additional equipment, sleep lab environment, and educated staff to reach features of prediction. The study of Aaronson et al. used self-reported symptoms, and socio-demographic and clinical features to build a prediction model of OSA using backward multivariable logistic regression [7].

This study population consisted of only patients with stroke and those with other comorbidities were not included. Also, the overall result of the present study was in the same line with that of the above mentioned study. Results show that the AUC of the Neural Networks, Naïve Bayes, Logistic Regression, SVM and Random Forest models were above 0.70, which was good enough to discriminate the power of the models [10, 11]. The retrospective nature of data gathering was a limitation of this study. Thus, a further prospective study is needed to confirm the clinical use of the proposed prognosticators. In addition, only people who performed PSG in Nemazi hospital sleep lab were included. With respect to this issue, a national or cohort study should be planned.

5. Conclusions

According to the results, the use of machine learning method with self-reported symptoms and easily available features for modeling has a good efficiency to predict OSA. To screen high-risk OSA patients, physicians or medical staff might use it as a fast and cost-effective auxiliary tool.

Acknowledgements

We thank Dr. Samrad Mehrabi, the staff of sleep lab at Nemazi hospital and those who helped us in the design and implementation of this project. The authors wish to thank Mr. H. Argasi at the Research Consultation Center (RCC) of Shiraz University of Medical Sciences for his invaluable assistance in editing this manuscript.

References

[1] Senaratna CV, et al. Prevalence of obstructive sleep apnea in the general population: a systematic review Sleep medicine reviews 2017;34: 70-81.
[2] Ibáñez V, Silva J, Cauli O. A survey on sleep assessment methods. PeerJ 2018;6: e4849.
[3] Collop NA, et al. Clinical guidelines for the use of unattended portable monitors in the diagnosis of obstructive sleep apnea in adult patients. Journal of clinical sleep medicine 2007;3: 737-747.
[4] Roebuck A, et al. A review of signals used in sleep analysis. Physiological measurement 2013;35: R1.
[5] Classification techniques on computerized systems to predict and/or to detect Apnea: A systematic review. Computer methods and programs in biomedicine 2017;140: 265-274.
[6] Wu MF, 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 journal of biomedical and health informatics 2016;21: 1524-1532.
[7] Aaronson JA, 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?, Archives of physical medicine and rehabilitation 2014;95, 747-752.
[8] Liu WT, Wu HT, Juang JN, Wisniewski A, Lee HC, Wu D, Lo YL. Prediction of the severity of obstructive sleep apnea by anthropometric features via support vector machine. PloS one 2017;12.
[9] Can we avoid unnecessary polysomnographies in the diagnosis of obstructive sleep apnea? A Bayesian network decision support tool, in: 2014 IEEE 27th International Symposium on Computer-Based Medical Systems, IEEE 2014, p. 28-33.
[10] Altman DG, Vergouwe Y, Royston P, Moons KG. Prognosis and prognostic research: validating a prognostic model. Bmj 2009;338: b605.
[11] Šimundić AM. Measures of diagnostic accuracy: basic definitions. Ejifcc 2009;19: 203.