Determination of Sleep Apnea Severity Using Multi-Layer Perceptron Neural Network

Zeinab Kohzadi, MSc¹, Reza Safdari, PhD², Khosro Sadeghniai Haghighi, MD, PhD³

¹Medical Informatics, Health Information Management Department, School of Allied Medical Sciences, Tehran University of Medical Sciences, Tehran, Iran

²Health Information Management Department, School of Allied, Medical Sciences, Tehran University of Medical Sciences, Tehran, Iran

³Occupational Sleep Research Center, Tehran University of Medical Sciences, Tehran, Iran

Background and Objective Sleep apnea is a rather common illness, which occurs due to dyspnea during night sleep. The effects of this illness can cause problems in the patient's life and affect its quality. Therefore, its timely diagnosis, using machine algorithms can be an important step towards preventing and controlling this illness.

Methods In this study is using artificial neural networks, in order to detect the severity of sleep apnea among 200 patients, who visited the Imam Khomeini sleep clinic in Tehran. Then the artificial neural network with the structure (8-10-3-1), Sigmoid transfer function and 120 educational cycles were designed and educated based on 70% of the data at hand. The artificial neural network was designed, using MATLAB2018.

Results Using the multi-layer perceptron classifier with 10-fold cross validation tests led to 96.5%, 92.4%, 91.5% and 94.5% correctness, respectively for normal, mild, moderate and severe classifications. Enough correctness of the algorithm reduces the patients' need to take the polysomnography test.

Conclusions The results show that using artificial neural network can be useful in detecting the sleep apnea severity, without using costly tests and limited PSG.

Sleep Med Res 2020;11(2):70-76

Key Words Sleep apnea, Neural network, Polysomnography, Wrapper, MLP.

Received: August 23, 2020

Revised: September 29, 2020

Accepted: November 6, 2020

Correspondence

Reza Safdari, PhD

Health Information Management Department,
School of Allied, Medical Sciences,
Tehran University of Medical Sciences,
Tehran, Iran

Tel +989183404581

Fax +88983037

E-mail Rsafdari@tums.ac.ir

ORCID

Zeinab Kohzadi

<https://orcid.org/0000-0001-8038-5454>

Reza Safdari

<https://orcid.org/0000-0002-4982-337X>

Khosro Sadeghniai Haghighi

<https://orcid.org/0000-0001-5242-5113>

© 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.

INTRODUCTION

Sleep is an inseparable part of human life, and any person spends one third of their lives sleeping. According to medical researches, there are 80 types of sleep disorders, for example, sleeping late, waking up too many times during the sleep, waking up early, having difficulty sleeping an sleep apnea [1-3]. But one of the most important disorders that can be dangerous, and in some cases even deadly, is sleep apnea, and approximately 2% to 4% of adults suffer from it. Apnea is the lack of air flow in nose and mouth for at least 10 seconds and each 25% to 50% reduction in air flow, during breathing, which occurs along with severe saturated oxygen drop in blood, is called hypopnea [4,5]. In long term, this disorder leads to sleepiness during the day [6], depression [7], lowered daily performance and quality of life [8], increased risk of accidents and incidents during driving and working [9], cardiovascular diseases, stroke and diabetes [10]. The standard way to diagnose sleep apnea is polysomnography (PSG), this test includes direct observation of the patient, along with electroencephalogram (EEG) control hypertension, breathing rhythm, heart rhythm, oxygen saturation, eye movements and muscles' electric actions. This test is used to discriminate central, obstructive or mixed apnea and calculates the apnea-hypopnea index by dividing the sum of apneas by the hours of sleep [11]. However, this device is too costly and its interpretation requires experts and it is not available everywhere [12].

Most of the researches, conducted in the field of diagnosing apnea, have used vital signals. In order to detect the severity of apnea, Gutierrez-Tobal et al. [13] examined the saturated blood oxygen signal, using pulse oximeter, of 320 patients at their homes. In 2018, 100 night records of single-channel air, based on relative entropy were used to automatically diagnose apnea-hypopnea events [14]. In another study, 70 record signals of single-lead EEG physionet database, through regression model method, were used [15]. However, the smart evaluations using clinical data is too little in this field. In 2015, an algorithm, based on artificial neural network (ANN), using four input variables and two outputs, using the information from 201 patients, was presented [16]. Utilizing machine-learning methods, Bozkurt et al. [17] classified obstructive sleep apnea (OSA) severity, using three categories of variables (clinical data, symptoms and body examination).

The purpose of the present study is to present an algorithm, based on ANN, for detecting the severity of sleep apnea in the patients, compared to PSG test, so that in case of having a strong diagnosis capability, patient's sleep apnea conditions can be diagnosed, before undergoing PSG test, using models and clinical and demographic data [age, gender, body mass index (BMI), neck circumference, snoring, hypertension, smoking and Epworth Sleepiness Scale (ESS)]. This result is important in that it can both prevent effects and probable damages of PSG in patients, who do not need it, and avoid the diagnostic tests costs.

METHODS

This study is of descriptive-analytic type, which has been conducted in five stages include dataset, preprocessing, variable selection, model training and test, classification result.

Dataset

In this study, 200 cases, including 134 males and 66 females, of patients, visiting Imam Khomeini sleep clinic in Tehran from October until November 2019, were examined. According to the expert practitioner, among these patients, 24 were in normal, 49 in mild, 91 in moderate and 36 in severe conditions. Demographic information, diagnostic tests and PSG results were among extracted variables from the patients' medical profiles. The data was determined, based on the sleep experts' opinions and also different studies. This data was recorded and analyzed in Excel (Microsoft, Redmond, WA, USA) and SPSS 2016 (IBM Corp., Armonk, NY, USA), in the form unassignable data. Independent variables in this study include sex, age, BMI, neck circumference, snoring, smoking, hypertension and ESS. Input variables were applied to the neural network, according to Table 1. In order to analyze the multi-layer perceptron (MLP) neural network with the LM algorithm, MATLAB 2018 (MathWorks, Natick, MA, USA) was used.

Preprocessing

Preprocessing techniques and data clearing are executed with the purpose of improving the quality. Existence of missing data in medical sciences is inevitable. Since in our dataset, there are features with missing values (12 cases) and we don't want to lose this information, therefore, before modelling, we need to fill these values with appropriate values, so the missing values are estimated by expectation-maximization (EM) algorithm. EM is one of the modern and advanced methods of solving the problem of missing data, which, theoretically, has certain complications, however, application-wise, has better performance, compared to classic methods, including questions average, individual average, individual mode and regression [18]. This algorithm is considered an effective repeating process for calculating maximum likelihood in the presence of missing data. Each repeat of the algorithm includes two steps: expected value step (step-E) and maximization step (step-M). Since the value of likelihood increases with each repeat of the algorithm, we can be sure of convergence [19,20].

Variable Selection

Wrapper method has been used for choosing variables in this study. The process is shown in Fig. 1.

In this method, induction algorithm is used for choosing features. All of the variables' states are put into the model and considered. There are 2^n possibilities for n variables [21]. Wrapper method requires high volumes of calculation. For example in this

Table 1. Sleep apnea risk factors

Variables	Description
Sex	
Male	1
Female	0
Age	100
BMI	100
Neck circumference	100
Hypertension	
Yes	1
No	0
Smoking	
Present	1
Absent	0
Snoring	
Yes	1
No	0
ESS	
> 10	1
< 10	0

BMI: body mass index, ESS: Epworth Sleepiness Scale.

study, for 8 variables, $2^8 = 256$ states were examined. This method is strong against over fitting [22].

Model Training and Test

Artificial MLP neural network, due to its valuable parallel capabilities and learning, is mostly used for solving complex problems. The general model of perceptron networks is a back-propagating progressive network. Progressive networks are networks, the first neuron layer inputs of which are connected to the next layers, and this is true in each level until it reaches the output layer. Back-propagation process means that after the determination of the network's output, first the last layers' weights are corrected and then the previous layers' weights will be corrected [23].

In Fig. 2, the MLP network's structure is shown. It is assumed that there are M layers, and in each layer there are J_m nodes. The "m-1" layer's connection weight to the "m" layer is shown as $W^{(m-1)}$. Also bias, output and forcing function, "I" neuron from the "m" layer are, respectively, introduced by $o_i^{(m)}$, $\Phi_i^{(m)}$, and $\theta_i^{(m)}$ [24].

The output of all processing units, from each layer, is given to all the processing units of the next layer. Processing units in the

layer are all linear, but in hidden layers, specially output layer, non-linear neurons, with any other non-linear continuous and differentiable function can be used. The forcing function used in MLP in this study is Sigmoid function [25].

Learning algorithm

Neural networks are capable of learning from the past, experience and environment, and improve their own behavior while doing so. In order to educate the MLP neural network, learning with observer method is used [26]. In the designed neural network, Levenberg-Marquardt algorithm is used, because of its faster convergence in educating average-size networks [27].

Implementing a neural network consists of three parts: preparing test samples, education phase and testing the neural network.

Neural network education phase

In designing the aforementioned neural network, two matrices, the first one with 200 samples and 8 features and the second one with 200 samples and four Status (1, 2, 3, and 4), which respectively indicate normal, mild apnea, moderate apnea and severe apnea, were used as objective matrices. 70% of this preprocessed data was used for educating the network.

Testing the neural network phase

In this part, the 30% of the preprocessed data (15% for validation and 15% for testing) that was not used in the education phase, is implemented into the ANN in the form of a matrix, and put into the software. In order to examine the network's success and performance, three factors, precision, sensitivity and feature, of the confusion matrix were used, through the following equations. In order to achieve more precision in the evaluation of the final results, the test was repeated 10 times in each design, on average.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

Fig. 1. Choosing variables by wrapper method.

Fig. 2. Multilayer perceptron neural network structure.

$$\text{Sensitivity} = \frac{TP}{TP + FN} \quad (2)$$

$$\text{Specificity} = \frac{TN}{TN + FP} \quad (3)$$

Table 2. Descriptive statistics of quantitative variables

Variable	SD $\pm$ $\bar{X}$
Age	42.3 $\pm$ 11.2
BMI	26.6 $\pm$ 5.2
Neck circumference	38.2 $\pm$ 4.8
ESS	8.2 $\pm$ 6.5

$\bar{X}$: sample mean, SD: standard deviation, BMI: body mass index, ESS: Epworth Sleepiness Scale.

Table 3. Descriptive statistics of qualitative values

Variable	Percentage
Sex	
Male	64
Female	36
Hypertension	
Yes	65
No	35
Smoking	
Yes	52
No	48
Snoring	
Present	76
Absent	24

Table 4. Comparison of MLP neural network's diagnostic power for different variables

Best network structure	Variable selection	Average accuracy	Variable count
1-5-2-1	Sex	64.1	1
1-6-3-4	Age	69	1
1-5-2-1	Hypertension	69.7	1
1-6-6-1	BMI	70.3	1
1-7-1-1	Neck circumference	70.1	1
1-4-5-1	Smoking	61.5	1
1-9-3-1	Snoring	76	1
1-8-2-1	ESS	75.2	1
2-5-1-1	ESS, snoring	80.1	2
3-3-7-1	Snoring, ESS, BMI	83.7	3
4-6-2-1	Snoring, ESS, BMI, age	85.2	4
5-3-3-1	Snoring, ESS, BMI, age, sex	87.1	5
6-1-5-1	Snoring, ESS, BMI, age, sex, neck circumference	89	6
7-3-6-1	Snoring, ESS, BMI, age, sex, neck circumference, hypertension	90.1	7
8-10-3-1	Snoring, ESS, BMI, age, sex, circumference, smoking, hypertension	93.62	8

MLP: multi-layer perceptron, BMI: body mass index, ESS: Epworth Sleepiness Scale.

TP: true positive, TN: true negative, FP: false positive, FN: false negative.

Ethical Approval

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Informed consent was obtained from all individual participants included in the study.

RESULTS

In order to create a general view, variables were divided into two quantitative and qualitative categories. Tables 2 and 3 show the statistics, related to these variables.

In order to find the missing data, EM algorithm was used in this study. And for choosing variables, metaheuristic optimization algorithm Wrapper was used. Table 4 shows the diagnostic power of the MLP perceptron neural network's architecture for different layers and also the achieved correctness average and the number of chosen variables.

According to Table 4, the best architecture for MLP neural network is (8-10-3-1), which has 8 inputs, 2 hidden layers, one with 10 neurons and the other with 3 neurons and 1 outputs. The execution of this network on MATLAB is shown in Fig. 3.

The aforementioned neural network's education lasted 1 minute and 36 seconds, and after 6 repeats, i.e. from 121 to 126, due to lack of improvement in the network's education process, this process was stopped. The least mean square (0.07) was achieved

Fig. 3. Neural network's execution.

Table 5. Calculating the values of accuracy, specificity, and sensitivity

	Normal	Mild	Moderate	Severe
Accuracy (%)	96.5	92.4	91.5	94.5
Specificity (%)	97.7	94.7	92.6	95.3
Sensitivity (%)	87.5	83.6	90.1	86.1

in epoch 372.

The designed neural network correctly diagnosed 21 cases of normal apnea, 41 cases of mild apnea, 82 cases of moderate apnea and 31 cases of severe apnea. The correctness, feature and sensibility of the aforementioned network are presented in Table 5.

DISCUSSION

Sleep apnea is a disorder, which is practically a serious danger towards life, and in some cases, it needs fast clinical interventions. Given the numerous effects this illness causes for people, the importance of its timely treatment increases daily and attracts more attention. Effective treatment of sleep apnea requires to detection its severity [28].

The common problem in automatic medical diagnosis is finding the best and fastest possible algorithm, which does not require much time and yields the best results. In order for this to be realized, there needs to be a strong and reliable medical diagnosis system that supports the complicated diagnosis process and minimizes the possible mistakes, made by experts [29]. Therefore, the present study was conducted with the purpose of presenting a model to help medical diagnosis in detecting the severity of apnea, using MLP ANN. The ANN is a non-parametric method for classification. In the neural network structure, first, the network was educated by educational samples and Levenberg-Marquardt algorithm, and was tested by test samples, then after examining all the possible models in MLP, the best model, i.e. NN (8-10-3-1), with the presence of gender, age, BMI, smoking, neck circumference, hypertension and ESS, was achieved, since the least square mean was small and its correctness in diagnosing normal, mild, moderate and severe was 96.5%, 92.4%, 91.5%, and 94.5% respectively. The results show that the designed neural network has been successful in diagnosing the apnea se-

verity, and has completed the classification with proper precision.

There are not many studies on detecting the apnea severity, using clinical data and artificial intelligence algorithms.

Viner et al. [30] used logistic regression model for diagnosis. They examined a group of 410 people, the rate of apnea among which was 46%. And the important variables, achieved by this are age, gender, snoring, BMI. Sensitivity, feature and effectiveness (the area under the ROC curve) of the proposed model are, respectively, 94%, 28% and 77%.

Karmanli et al. [16] examined 201 cases of patients, including 140 OSA patients and 61 healthy persons. They used four variables (sex, age, and BMI) as neural network's inputs and the outcome was yes or no. the achieved correctness, using MLP classifier with 20 neurons in the middle layer, was 86.8%. The high number of neurons in the middle layer may cause ANN some problems, such as prolonging the network's education test time, and also the network may learn the unimportant system of the educational data and perform weakly in solving problems. In our study, we used 12 neurons in the middle layer and more input variables, therefore, higher correctness for apnea-hypopnea index, in all conditions, was achieved.

Bozkurt et al. [17] attempted to diagnose the severity of apnea, using three categories of variables (clinical data, symptoms and physical diagnosis) and Decision tree, Bayesian network, Random forest, Neural Network and Logistic regression classifiers on cases. Each model was educated and evaluated, using cross-validation ten times, and for evaluating the execution of classifying all methods, true positive rate (TPR), false positive rate (FPR), predicted positive value, F-measure and the area under the curve were used. The highest TPR was 0.71 and the lowest FPR was 0.15. In our study, using Wrapper method for choosing variables, a proper amount of variables was chosen and higher correctness was achieved.

The main advantage of ANNs is their non-linear and flexible modelling capability. These kinds of networks do not require recognizing the special form of the model, and the model is formed based on the information, available in the data. As previously mentioned in the article, ANNs, in addition to their vast area of application, are better tools for prediction and diagnosis, compared to statistical methods. This study shows the ANN's highly precise diagnostic capability in relation to data on sleep apnea. Moreover, it also verifies the other studies in the field of ANNs;

therefore, using ANNs in medical studies is recommended.

In order to continuously improve the results of applying ANN models, simultaneously using ANNs and pattern recognition, such as decision tree and fuzzy algorithms, for using the created rules and extracting features, will be useful. Working on activation functions, which leads to simpler structure and faster convergence speed in neural network models, and optimizing neural network, using evolutionary algorithms for weighting network and bigger database with more records in the next studies are some of the methods of increasing the implemented neural network models precision. Neural networks can also be used for diagnosing other illnesses, because, due to its low cost and fast speed of performance, it will be time and cost effective.

Conclusions

Since in medical sciences researches, human's health is at stake, correct diagnosis of the results is of great importance, therefore, methods, which yield the least errors and highest reliability in predicting and diagnosing, must be used. One of the methods that has attracted the attention of many researchers is ANN. In this study, by using clinical data and implementing ANN, high and acceptable values for correctness, sensitivity and feature, in detecting the severity of sleep apnea, have been achieved. The results from the sleep apnea severity detection model, using artificial intelligence and clinical data can be very important. Because in addition to reducing the costs of using PSG and preventing its possible harms and side effects, they also reduce possible errors and mistakes in diagnosis, made due to tiredness or inexperience of clinical experts, and recognizes the patients, who need these diagnostic and clinical measures, with the highest precision and in the fastest time. Implementing the sleep apnea severity detection algorithm, by designing mobile-based and web-based user-friendly user interface, or in the form of a soft ware, given its accessibility anywhere and anytime and its lower costs, compared to PSG, can be a suitable supplement for PSG.

Acknowledgments

The authors would like to gratefully acknowledge the contribution department of health information management and Imam Khomeini Hospital Sleep Clinic at Tehran University of Medical Sciences. This paper was developed as part of a MSc thesis that is funded and supported by Tehran University of Medical Sciences.

Conflicts of Interest

The authors have no financial conflicts of interest.

Authors' Contribution

Conceptualization: all authors. Data curation: Kohzadi Z. Formal analysis: Kohzadi Z. Investigation: all authors. Methodology: Kohzadi Z. Project administration: Safdari R. Resources: Safdari R, Haghighi KS. Software: Kohzadi Z, Safdari R. Supervision: Safdari R, Haghighi KS. Validation: Safdari R, Haghighi KS. Visualization: all authors. Writing—original draft: Kohzadi Z. Writing—review & editing: Kohzadi Z, Haghighi KS.

REFERENCES

- Hirotsu C, Albuquerque RG, Nogueira H, Hachul H, Bittencourt L, Tufik S, et al. The relationship between sleep apnea, metabolic dysfunction and inflammation: the gender influence. *Brain Behav Immun* 2017;59: 211-8.
- Chung KF. Use of the Epworth Sleepiness Scale in Chinese patients with obstructive sleep apnea and normal hospital employees. *J Psychosom Res* 2000;49:367-72.
- Whited MC, Olendzki E, Ma Y, Waring ME, Schneider KI, Appelhans BM, et al. Obstructive sleep apnea and weight loss treatment outcome among adults with metabolic syndrome. *Health Psychol* 2016;35:1316-9.
- Bhushan B, Ayub B, Thompson DM, Abdullah F, Billings KR. Impact of short sleep on metabolic variables in obese children with obstructive sleep apnea. *Laryngoscope* 2017;127:2176-81.
- Lurie A. Obstructive sleep apnea in adults: relationship with cardiovascular and metabolic disorders. *Adv Cardiol* 2011;49.
- Vaessen TJ, Overeem S, Sitskoorn MM. Cognitive complaints in obstructive sleep apnea. *Sleep Med Rev* 2015;19:51-8.
- Kerner NA, Roose SP. Obstructive sleep apnea is linked to depression and cognitive impairment: evidence and potential mechanisms. *Am J Geriatr Psychiatry* 2016;24:496-508.
- Teng Y, Xiong Y, Wang N. The applications of the STOP-Bang questionnaire in screening obstructive sleep apnea in patients with metabolic syndrome. *Zhonghua Jie He He Hu Xi Za Zhi* 2015;38:461-6.
- Jang HU, Park KS, Cheon SM, Lee HW, Kim SW, Lee SH, et al. Development of the Korean version of the sleep apnea quality of life index. *Clin Exp Otorhinolaryngol* 2014;7:24-9.
- Naegele B, Pepin JL, Levy P, Bonnet C, Pellat J, Feuerstein C. Cognitive executive dysfunction in patients with obstructive sleep apnea syndrome (OSAS) after CPAP treatment. *Sleep* 1998;21:392-7.
- Ciolek M, Niedźwiecki M, Sieklicki S, Drozdowski J, Siebert J. Automated detection of sleep apnea and hypopnea events based on robust airflow envelope tracking in the presence of breathing artifacts. *IEEE J Biomed Health Inform* 2015;19:418-29.
- Gutta S, Cheng Q, Nguyen HD, Benjamin BA. Cardiorespiratory model-based data-driven approach for sleep apnea detection. *IEEE J Biomed Health Inform* 2018;22:1036-45.
- Gutierrez-Tobal GC, Alvarez D, Crespo A, Del Campo F, Hornero R. Evaluation of machine-learning approaches to estimate sleep apnea severity from at-home oximetry recordings. *IEEE J Biomed Health Inform* 2019;23:882-92.
- Jia Z, Li J, Huang JJ, Chen L, Yang L, Zhang T. Automated diagnosis of the sleep apnea hypopnea syndrome based on adjusted relative entropy. *Chin J Ophthalmol Otorhinolaryngol* 2018;18:382-388.
- Wang L, Lin Y, Wang J. A RR interval based automated apnea detection approach using residual network. *Comput Methods Programs Biomed* 2019;176:93-104.
- Karamanli H, Yalcinoz T, Yalcinoz MA, Yalcinoz T. A prediction model based on artificial neural networks for the diagnosis of obstructive sleep apnea. *Sleep Breath* 2016;20:509-14.
- Bozkurt S, Bostanci A, Turhan M. Can statistical machine learning algorithms help for classification of obstructive sleep apnea severity to optimal utilization of polysomnography resources? *Methods Inf Med* 2017;56:308-18.
- Dempster AP, Laird NM, Rubin DB. Maximum likelihood from incomplete data via the EM algorithm. *J R Stat Soc Series B Stat Methodol* 1977;39:1-38.
- Ghomrawi HM, Mandl LA, Rutledge J, Alexiades MM, Mazumdar M. Is there a role for expectation maximization imputation in addressing missing data in research using WOMAC questionnaire? Comparison to the standard mean approach and a tutorial. *BMC Musculoskelet Disord* 2011;12:109.
- Downey RG, King C. Missing data in Likert ratings: a comparison of replacement methods. *J Gen Psychol* 1998;125:175-91.

21. Kohavi R, John GH. Wrappers for feature subset selection. *Artif Intell* 1997;97:273-324.
22. Guyon I, Elisseeff A. An introduction to variable and feature selection. *J Mach Learn Res* 2003;3:1157-82.
23. Cross SS, Harrison RF, Kennedy RL. Introduction to neural networks. *Lancet* 1995;346:1075-9.
24. Du KL, Swamy MN. *Neural networks in a softcomputing framework*. Germany: Springer Science and Business Media 2006.
25. Hagan MT, Demuth HB, Beale M. *Neural network design*. Boston, MA: Pws Pub. 1996.
26. Livingstone DJ. *Artificial neural networks: methods and applications*. Springer: Humana Press 2008.
27. Mobley BA, Schechter E, Moore WE, McKee PA, Eichner JE. Neural network predictions of significant coronary artery stenosis in men. *Artif Intell Med* 2005;34:151-61.
28. Sia CH, Hong Y, Tan LWL, van Dam RM, Lee CH, Tan A. Awareness and knowledge of obstructive sleep apnea among the general population. *Sleep Med* 2017;36:10-7.
29. Kumar K, Abhishek. Artificial neural networks for diagnosis of kidney stones disease. *IJITCS* 2014;4:20-25.
30. Viner S, Szalai JP, Hoffstein V. Are history and physical examination a good screening test for sleep apnea? *Ann Intern Med* 1991;115:356-9.