The Factors of Subjective Voice Disorder Using Integrated Method of Decision Tree and Multi-Layer Perceptron Artificial Neural Network Algorithm

Haewon Byeon
Department of Speech Language Pathology
Nambu University, Gwangju, Republic of Korea

Sunghyoun Cho*
Department of Physical Therapy
Nambu University, Gwangju, Republic of Korea

Abstract—The aim of the present study was to develop a prediction model for subjective voice disorders based on an artificial neural network algorithm and a decision tree using national statistical data. Subjects of analysis were 8,713 adults over the age of 19 (3,801 males and 4,912 females) who completed the otolaryngological examination of the Korea National Health and Nutrition Examination Survey from 2010 to 2012. Explanatory variables included age, education level, income, occupation, problem drinking, coffee consumption, and pain and discomfort from disease over the last two weeks. A multi-layer perceptron artificial neural network and a decision tree model were used for the analysis. In this model, smoking, pain and discomfort from disease over the last two weeks, education level, occupation, and income were drawn out as major predictors of subjective voice disorders. In order to minimize the risk of dysphonia, it is necessary to establish a scientific management system for high-risk groups.

Keywords—Neural Networks; Subjective Voice Disorder; decision tree; risk factor; data-mining

I. INTRODUCTION

Voice disorders refer to problems of the voice due to abnormalities in the structure, function, or vagus nerve of vocal cords, and the term comprises both laryngeal disorders and subjective voice problems [1]. The prevalence rate of voice disorders is 5–7% [2, 3], and 30% of the community-dwelling population experience voice problems at least once in their lifetime [4]. Thus, it is presumed that among the total population of 50 million Koreans, more than 2.5 million Koreans suffer from voice disorders.

The measurement and assessment of voice disorders are classified into objective methods, such as acoustic and aerodynamic tests, and patients' subjective assessments of voice symptoms [5]. Preliminary tests for the diagnosis of voice disorders include the acoustic test and the laryngoscopy test, but it is difficult to detect functional dysphonia caused by psychological problems with these tests alone. Therefore, capturing the subjective perceptions of a voice problem that a subject reports as well as observing the voice problems objectively is important in the diagnosis of dysphonia [6]. In addition, although there are (individual) differences in the degree of self-recognition of voice problems depending on the characteristics of individuals, subjective voice problems play an important role in deciding the treatment of dysphonia [7].

Nevertheless, in many cases, subjects do not take voice problems seriously and, thus, do not report them to medical professionals. Moreover, even when subjects do report, diagnoses are terminated if no problems are discovered on the objective tests.

In order to effectively prevent dysphonia, investigations of its risk factors are vital. Over the last 20 years, smoking, drinking, misuse and abuse of vocal cords, and occupation have been reported to be the major risk factors of dysphonia [8–12]. Among them, subjective voice problems have been verified to be not only an independent risk factor of voice disorders [3, 4, 13, 14] but also the most predominant predictor among the various risk factors of dysphonia [15].

To date, numerous epidemiological studies have confirmed that subjective voice disorders are an independent risk factor of dysphonia [16]. However, it has not yet been verified as a risk factor in Korean adults.

Meanwhile, artificial neural network analysis, an analysis technique that increases problem-solving abilities through the learning of artificial neurons, is currently being widely used in classification and prediction. Inspired by the human brain, neural network analysis has several advantages [17]. For example, past experiences enable automatic learning, the analysis of qualitative and quantitative variables is possible, and it has excellent prediction power, as non-linear combinations among entered variables are possible.

Furthermore, the decision tree model, which displays the decision-making process in a tree-structure diagram, has the advantages of accommodating both continuous and categorical variables and enabling an understanding of the factors contributing to the dependent variable.

In order to determine the characteristics of high-risk groups for dysphonia, it is necessary to elucidate the complex factors that affect voice problems. Decision trees, widely used in the areas of pattern recognition and medical science, and data mining analysis, used in neural networks, can be effectively used in predicting the target group for dysphonia prevention programs [15].

This study presents basic materials to reduce dysphonia by developing a prediction model for subjective voice disorders based on an artificial neural network algorithm and a decision tree using national statistical data. The organization of our

study is as follows: Section 2 describes the data resources, Section 3 explains the procedure for the development of the prediction model, and Section 4 presents the results of the developed prediction model.

II. METHODS

A. Data sources

This study analyzed 8,713 adults (3,801 males and 4,912 females) who completed the otolaryngological examination of the Korea National Health and Nutrition Examination Survey from 2010 to 2012. The Korea National Health and Nutrition Examination Survey is a nationwide health survey conducted by the Ministry of Health and Welfare on 11,520 households regarding education, economic activities, contraction of diseases, use of medical institutions, and health behaviors [18]. Education and economic activities were researched by individual face-to-face interviews, while health behaviors, such as smoking and drinking, were researched by self-administered questionnaires. The detailed methods of the research are specified in the preceding study [18].

B. Measurements

The dependent variable, subjective voice disorders, was classified (yes, no) based on the answers to the otolaryngological question: "Do you think that you have an abnormality in your voice?"

Explanatory variables included age, education level, income (quartile), occupation, problem drinking, coffee consumption, and pain and discomfort from disease over the last two weeks (yes, no). Age was classified as 19 to 39 years, 40 to 59 years, and more than 60 years old. Education levels were classified as below elementary school graduation, middle school graduation, high school graduation, and above college graduation. Occupations were classified as follows: economically inactive person, non-manual worker (e.g., managers & professionals, clerical support workers, service & sales workers), manual worker (skilled agricultural & forestry & fishery workers, craft & plant and machine operators and assemblers, and unskilled laborers). As for alcohol consumption, 8 points and over was classified as problem drinking by using Alcohol Use Disorders Identification Test in Korea (AUDIT-K)[19].

III. STATISTICAL ANALYSIS

A. Artificial neural network

Factors potentially related to subjective voice disorders were analyzed by using an artificial neural network. Artificial neural network analysis is a data mining modeling technique that finds hidden patterns from actual data through a repetitive learning process imitating the neural network of the human brain. It is a nonlinear model that is used to solve prediction problems in data with complex structures [20].

Artificial neural network analysis is a mathematical model composed of numerous processing factors with a hierarchical structure, and it learns the relationship between input and output by the repetitive adjustment of weights by comparing past input data values and corresponding output data values [21]. The structure of the neural network is composed of both

an input layer made of nodes corresponding with input variables and a hidden layer made of multiple hidden nodes. The hidden nodes turn the linear combination of variable values delivered from the input layer into a nonlinear function and deliver it to the input layers or other hidden layers [22].

This study used a Radial Basis Function (RBF) neural network [23]. This study regarded variables with relative importance of inputs over 0.1 as major explanatory variables that affect the dependent variable and thus included them in the decision tree model.

B. Classification and regression tree algorithm

The decision tree model was established by using the Classification and Regression Tree (CART) algorithm. CART is an algorithm based on binary classification that measures impurities by using the Gini Index, and in it, only two children nodes are formed from a parent node [24]. The Gini Index refers to the probability that two elements randomly extracted from n elements belong to different groups from each other [25]. The alpha value for the criteria of splitting and merging was set at 0.05. The number of parent nodes was 200 and that of child nodes was 100, and the number of branches was limited to five. The validity of the developed model was assessed with the 10-fold cross-validation method.

IV. RESULTS

A. General characteristics of study subjects

The general characteristics of the study subjects are presented in Table 1. Out of the total 8,713 subjects, the prevalence of subjective voice disorders was 6.9% (n=602). According to the result of the chi-square test, healthy subjects and those with subjective dysphonia did not have significant differences in any of the variables.

TABLE I. GENERAL CHARACTERISTICS OF THE STUDY SUBJECTS, N (%)

Characteristics	Subjective voice disorder		p
	No (n=8,111)	*Yes (n=602)*	
Age			0.79
19-39	2,485 (93.0)	188 (7.0)	
40-59	2,050 (93.3)	218 (6.7)	
60 ≤	2,576 (92.9)	196 (7.1)	
Sex			0.75
Male	3,542 (93.2)	259 (6.8)	
Female	4,569 (93.0)	343 (7.0)	
Education			0.31
Elementary school	1,885 (93.1)	139 (6.9)	
Middle school	835 (92.7)	66 (7.3)	

High school	2,444 (92.7)	192 (7.3)	
Collage	2,267 (93.9)	146 (6.1)	
Income			0.35
First quartile	1,579 (93.5)	109 (6.5)	
Second quartile	2114 (92.9)	162 (7.1)	
Third quartile	2,138 (92.4)	177 (7.6)	
Fourth quartile	2,150 (93.6)	148 (6.4)	
Occupation			0.46
economically inactive person	3,067 (93.3)	221 (6.7)	
non-manual worker	2,470 (92.8)	193 (7.2)	
manual worker	1,874 (93.7)	127 (6.3)	
Problem drinking			0.91
No	4,263 (93.1)	317 (6.9)	
Yes	2,095 (93.2)	154 (6.8)	
Smoking			0.53
Non-smoking	4,323 (93.4)	306 (6.6)	
Past smoking	1,547 (92.6)	124 (7.4)	
Current smoking	1,561 (93.2)	114 (6.8)	
Pain & discomfort from disease for the recent 2 weeks			0.59
No	5,734 (93.3)	412 (6.7)	
Yes	1,704 (92.8)	133 (7.2)	

B. Factors potentially related to subjective voice disorders

As a result of the artificial neural network analysis on 60.6% of the training sample, 29.6% of the test sample and 9.8% of the verification sample, five hidden layers were drawn out that produced the smallest data errors. The sum of the square error was 7.2%, and the classification accuracy of the training sample, test sample, and verification sample proved to be 92.8%, 93.8%, and 92.6%, respectively.

A synaptic weighted network diagram of the neural network model is presented in Figure 1. The synaptic weighted value in the network diagram demonstrates the relationship among layers, and the higher the combined weighted value, the thicker the line between layers. In this model, smoking, pain and discomfort from disease over the last two weeks, education level, occupation, and income were drawn out as major variables with high weighted values for subjective voice disorders.

The normalized importance sampling estimator drawn out from the neural network model is presented in Figure 2. According to the result of the normalized importance sampling estimator, smoking, pain and discomfort from disease over the last two weeks, education level, occupation, and income were deciding factors of subjective voice disorders.

The prediction model for subjective voice disorders using the CART algorithm is presented in Figure 3. According to the result of the classification model constructed using the CART algorithm, the most preferentially involved predictor was household income.
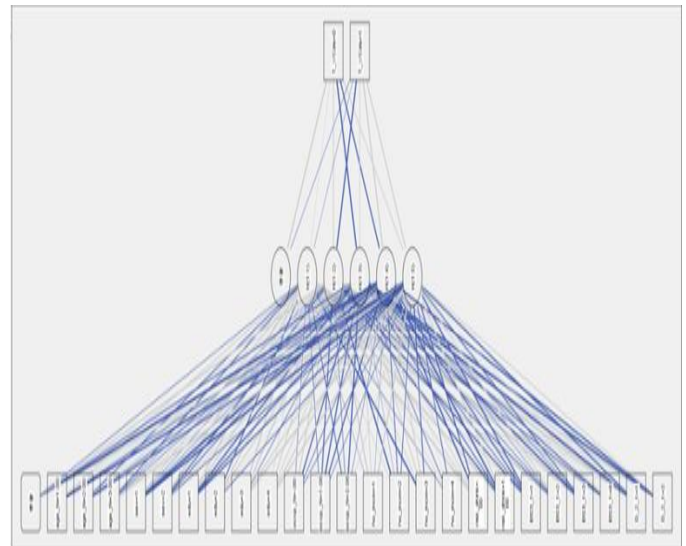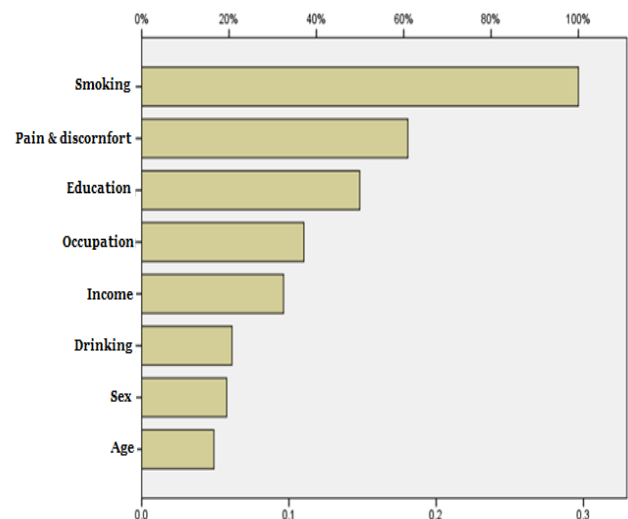

Fig. 1. Synaptic weighted network diagram

Fig. 2. The normalized importance sampling estimator drawn out from the neural network model

Table 2 is a gains chart of the final prediction model for subjective voice disorders created by the CART algorithm. The two nodes were confirmed as significant paths that effectively predicted subjective voice disorders.

The first path with the greatest profit index value (10th node) was "past smokers who belong to the lowest 25% in household income and are middle school graduates and lower or university graduates and over." Of them, 13.5% were classified as suffering from subjective dysphonia, and their profit index was 204.8%. The second path (16th node) was "manual workers who belong to the lowest 25% in household income and are primary school graduates and lower or high school graduates with past smoking experience." Of them, 11.1% were classified as suffering from subjective dysphonia, and their profit index was 169.1%.
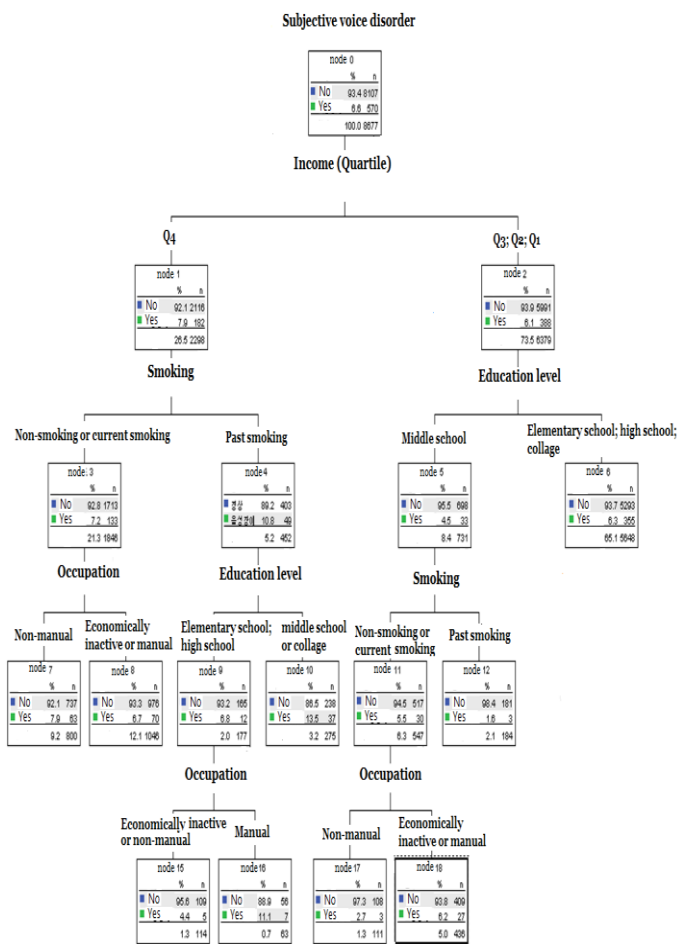

Fig. 3. The prediction model for subjective voice disorders using the CART algorithm

This study conducted a 10-fold cross-validation test to assess the final developed model. According to the result of the 10-fold cross-validation test to compare the stability of the drawn-out model, the drawn-out risk index was 0.32 and the misclassification rate was 32% for the cross-classification model, showing the same risk index (0.32) and misclassification rate (32%) of the prediction model.

TABLE II. GAINS CHART OF THE FINAL PREDICTION MODEL FOR SUBJECTIVE VOICE DISORDERS CREATED BY THE CART ALGORITHM

Node no	Node n (%)[1]	Gain n (%)[2]	Response %[3]	Gain Index %[4]	Group
10	275 (3.2)	37 (6.5)	13.5	204.8	Past smokers who belong to the lowest 25% in household income and are middle school graduates and lower or university graduates and over
16	63 (0.7)	7 (1.2)	11.1	169.1	Manual workers who belong to the lowest 25% in household income and are primary school graduates and lower or high school graduates with past smoking experience

[1] Node n(%); node number, % to 8,713
[2] Gain n(%); gain number, % to 602
[3] Response (%): The fraction of the subjective voice disorder
[4] Gain index (%):= 204.8 in total 7 node

V. CONCLUSION

This study analyzed predictors of subjective dysphonia by combining neural network analysis and a decision tree model.

In our study, "past smokers who belong to the lowest 25% in household income and are middle school graduates and lower or university graduates and over" and "manual workers who belong to the lowest 25% in household income and are primary school graduates and lower or high school graduates with past smoking experience" were high-risk groups for subjective dysphonia.

In order to prevent dysphonia effectively, it is necessary to develop prevention programs that consider the characteristics of high-risk groups.

REFERENCES

[1] D. R. Boone, S. C. McFarlane, S. L. Von Berg, and R. I. Zraick, The voice and voice therapy. Boston, Allyn & Bacon Communication Sciences and Disorders, 2013.

[2] H. Byeon, Gender differences in risk factors of benign vocal fold disease in Korea: the fifth Korea National Health and Nutritional Examination Survey. Logoped Phoniatr Vocol, E-pub, doi:10.3109/14015439.2015.1004365, 2015.

[3] N. Roy, R. M. Merrill, S. D. Gray, and E. M. Smith, Voice disorders in the general population: prevalence, risk factors, and occupational impact. Laryngoscope, vol. 115, no. 11, pp. 1988-1995, 2005.

[4]   S. M. Cohen, Self-reported impact of dysphonia in a primary care population: an epidemiological study. Laryngoscope, vol. 120, no. 10, pp. 2022-2032, 2010.

[5]   A. E. Aronson, and D. Bless, Clinical voice disorders. New York, Thieme Medical Publishers, 2011.

[6]   E. Vilkman, Voice problems at work: a challenge for occupational safety and health arrangement. Folia Phoniatrica et Logopaedica, vol. 52, no. 1-3, pp. 120-125, 2000.

[7]   A. Behrman, L. Sulica, and T. He, Factors predicting patient perception of dysphonia caused by benign vocal fold lesions. Laryngoscope, vol. 114, no. 10, pp. 1693-1700, 2004.

[8]   P. N. Carding, S. Roulstone, K. Northstone, and ALSPAC Study Team, The prevalence of childhood dysphonia: a cross-sectional study. Journal of Voice, vol. 20, no. 4, pp. 623-630, 2006.

[9]   N. R. Williams, Occupational groups at risk of voice disorders: a review of the literature. Occupational Medicine, vol. 53, no. 7, pp. 456-460, 2003.

[10]   S. L. Thibeault, R. M. Merrill, N. Roy, S. D. Gray, and E. M. Smith, Occupational risk factors associated with voice disorders among teachers. Annals of Epidemiology, vol. 14, no. 10, pp. 786-792, 2004.

[11]   H. Byeon, Relationships among smoking, organic, and functional voice disorders in Korean general population. Journal of Voice, vol. 29, no. 3, pp. 312-316, 2015.

[12]   H. Byeon, A population-based cross-sectional study of alcohol consumption and risk of benign laryngeal disease in Korean adults. Journal of Voice, E-Pub: doi:10.1016/j.jvoice.2014.10.014, 2016.

[13]   H. Byeon, Prevalence of perceived dysphonia and its correlation with the prevalence of clinically diagnosed laryngeal disorders: the Korea national health and nutrition examination surveys 2010-2012. Annals of Otology, Rhinology & Laryngology, vol. 124, no. 10, pp. 770-776, 2015.

[14]   N. Roy, J. Stemple, R. M. Merrill, and L. Thomas, Epidemiology of voice disorders in the elderly: preliminary findings, Laryngoscope, vol. 117, no. 4, pp. 628-633, 2007.

[15]   H. Byeon, The risk factors of laryngeal pathology in Korean adults using a decision tree model. Journal of Voice, vol. 29, no. 1, pp. 59-64, 2015.

[16]   J. H. Hah, S. Sim, S. Y. An, M. W. Sung, and H. G. Choi, Evaluation of the prevalence of and factors associated with laryngeal diseases among the general population. Laryngoscope, vol. 125, no. 11, pp. 2536-2542, 2015.

[17]   R. C. Eberhart, Neural network PC tools: a practical guide. San Diego, Academic Press, 2014.

[18]   Ministry of Health and Welfare, Korea National Health and Nutrition Examination Survey 2010-2012. Seoul, Ministry of Health and Welfare, 2014.

[19]   J. S. Kim, M. K. Oh, B. K. Park, M. K. Lee, and G. J. Kim, Screening criteria of alcoholism by alcohol use disorders identification test(AUDIT) in Korea. Journal of the Korean Academy of Family Medicine, vol. 20, no. 9, pp. 1152-1159, 1999.

[20]   V. Yashchenko, Multidimensional neural-like growing networks: a new type of neural network. International Journal of Advanced Computer Science and Applications, vol. 6, no. 4, pp. 1-10, 2015.

[21]   M. K. Luka, I. A. Frank, and G. Onwodi, Neural network based Hausa language speech recognition. International Journal of Advanced Research in Artificial Intelligence, vol. 1, no. 2, pp. 39-44, 2012.

[22]   A. G. Eldin, A data mining approach for the prediction of hepatitis C virus protease cleavage sites. International Journal of Advanced Computer Science and Applications, vol. 2, no. 12, pp. 179-182, 2011.

[23]   M. A. Jayaram, G. K. Prashanth, and M. Anusha, Indexing of ears using radial basis function neural network for personal identification. International Journal of Advanced Computer Science and Applications, vol. 6, no. 7, pp. 28-33, 2015.

[24]   H. Byeon, Development of prediction model for endocrine disorders in the Korean elderly using CART algorithm. International Journal of Advanced Computer Science and Applications, vol. 6, no. 9, pp. 125-129, 2015.

[25]   H. Byeon, A prediction model for mild cognitive impairment using random forests. International Journal of Advanced Computer Science and Applications, vol. 6, no. 12, pp. 8-12, 2015.