

Is Deep Learning Better than Machine Learning to Predict Benign Laryngeal Disorders?

Haewon Byeon

Department of Medical Big Data
College of AI Convergence, Inje University
Gimhae 50834, Gyeongsangnamdo, South Korea

Abstract—It is important in otolaryngology to accurately understand the etiology of a laryngeal disorder, diagnose it early, and provide appropriate treatment accordingly. The objectives of this study were to develop models for predicting benign laryngeal mucosal disorders based on deep learning, naive Bayes model, generalized linear model, a Classification and Regression Tree (CART), and random forest using laryngeal mucosal disorder data obtained from a national survey and confirm the best classifier for predicting benign laryngeal mucosal disorders by comparing the prediction performance and runtime of the developed models. This study analyzed 626 subjects (313 people with a laryngeal disorder and 313 people without a laryngeal disorder). In this study, deep learning was the best model with the highest accuracy (0.84). However, the runtime of deep learning was 39min 41sec, which was a 10 times longer development time than CART (3min 7sec). This model confirmed that subjective voice problem recognition, pain and discomfort in the last two weeks, education level, occupation, mean monthly household income, high-risk drinker, and current smoker were major variables with high weight for the benign laryngeal mucosal disorders of Korean adults. Among them, subjective voice problem recognition was the most important factor with the highest weight. The results of this study implied that the prediction performance of deep learning could be better than that of machine learning for structured data, such as health behavior and demographic factors as well as video and image data.

Keywords—Benign laryngeal mucosal disorder; voice disorder; deep learning; Naive Bayes model; generalized linear model

I. INTRODUCTION

Laryngeal disorders include organic dysphonia, caused by the structural changes (anatomical changes) of the larynx including the vocal cords, and functional dysphonia, which changes voice due to health risk behaviors (e.g., smoking or drinking) and improper habits (e.g., abuse or misuse of voice) [1]. In particular, benign laryngeal disorders refer to laryngeal disorders except for laryngeal cancer, a malignant tumor [2]. They are caused by abnormalities in the nervous system, mucous membranes, and cartilage [3], and they are frequently found in the adult population [4]. Benign laryngeal disorders include vocal polyp, vocal nodule, vocal cyst, Reinke's edema, vocal sulcus, vocal scar, contact granuloma, and laryngeal papilloma [5][6].

The prevalence of laryngeal disorders was 6.6% based on the American population [7]. Roy et al. (2005) reported that at least 1 in 10 Americans had experienced voice problems at

least once in their lifetime [7]. There is not enough data regarding the prevalence of laryngeal disorders in South Korea. The Otolaryngology Examination Survey of the 2012 Korean National Health and Nutrition Survey reported that the prevalence of benign laryngeal disorders was approximately 2.5% in South Korea [8]. It was reported that the prevalence of laryngeal orders is higher among men than women and smokers than nonsmokers [9][10]. It was also reported that the risk of laryngeal disorders was 1.4 to 1.6 times higher in managers, professionals, and service & sales workers than economically inactive people [11][12].

Voice is a very critical function for maintaining daily life. Particularly, it is directly related to living for certain occupations such as teachers, announcers, and singers. Consequently, discovering a laryngeal disorder early for maintaining a healthy voice can greatly improve the quality of patients' life [13,14,15]. Therefore, it is important in otolaryngology to accurately understand the etiology of a laryngeal disorder, diagnose it early, and provide appropriate treatment accordingly.

To date, the most common risk factors causing benign laryngeal mucosal disorders are voice abuse and wrong vocalization habit [16,17,18,19,20,21,22]. Other very diverse factors (e.g., smoking, drinking, viral infection, upper respiratory tract infection, and laryngopharyngeal reflux) have also been reported as risk factors [16,17,18,19,20,21,22]. However, since a disease is a result of complex interactions between multiple risk factors, not caused by a single risk factor, it is limited to predict a disease by exploring only individual risk factors [23]. To make it harder, different treatments need to be given according to individual characteristics (habits) and etiology, even though the shape of the lesions of a laryngeal disorder on the vocal cord mucosa is similar [24]. Consequently, it is important to fully understand the etiology of a benign laryngeal mucosal disorder and identify multiple risk factors of the disease in order to perform accurate diagnosis and treatment. Nevertheless, most studies that have evaluated the risk factors of laryngeal disorders have just tried to find individual risk factors using regression analysis [25,26,27,28,29], and only a few studies have explored the multiple risk factors of benign laryngeal mucosal disorders using machine learning [30].

Supervised learning-based machine learning has been used as a way to detect a disease and identify multiple risks in recent years [31,32,33]. Many recent studies [34,35] have reported that neural network-based deep learning is more accurate in

classifying and predicting diseases than machine learning. Nevertheless, previous studies [36, 37] mainly focused on developing classifiers for discriminating the presence of laryngeal diseases by mostly using video and image data. However, there are not enough studies on developing models to predict benign laryngeal mucosal disorders while reflecting various features (e.g., health behavior, disease, and demographic characteristics) in health surveys. The objectives of this study were to develop models for predicting benign laryngeal mucosal disorders based on deep learning, naive Bayes model, generalized linear model, a Classification and Regression Tree (CART), and random forest using laryngeal mucosal disorder data obtained from a national survey and confirm the best classifier for predicting benign laryngeal mucosal disorders by comparing the prediction performance and runtime of the developed models.

Construction of this study is as follows: Section II explains data source, measurements, development and validation of prediction models. Section III compares the results of developed machine learning models. Lastly, Section IV presents conclusion and direction for future studies.

II. MATERIALS AND METHODS

A. Data Source

This study targeted adults (≥ 19 years old) who participated in the otolaryngology examination and completed the 2012 KNHANES. The KNHANES extracts survey plots using the proportional allocation systematic sampling method that stratifies administrative districts and types of residences across the country and extracts samples proportional to the population survey plots of each layer. This study selected 4,528 adults (313 subjects with a laryngeal disorder and 4,215 subjects without a laryngeal disorder) who completed the health questionnaire, the otolaryngology questionnaire, and laryngeal endoscopy as the primary subjects of this study. Since the prevalence of a laryngeal disorder was only 6.9% among the subjects, showing a data imbalance issue, this study resolved the imbalance issue by using propensity score matching, which matched sex and age (1:1 ratio). Finally, this study analyzed 626 subjects (313 people with a laryngeal disorder and 313 people without a laryngeal disorder).

This study conducted a power test using G-Power program 3.1.9 (Universität Mannheim, Mannheim, Germany) for the final analysis data. When power (1-B) was 0.95, significance level (alpha) was 0.05, effect size (f2) was 0.35, and 201 predictor variables were applied, the appropriate sample size was 361. Therefore, the sample size of this study (626) satisfied the appropriate sample size for testing statistical significance (Fig. 1).

B. Variables

Benign laryngeal disease [20] in this study were defined as vocal nodules, laryngeal polyps, intracordal cysts, reinke's edema, laryngeal granuloma, glottic sulcus and laryngeal keratosis (Fig. 2). The explanatory variables were occupation (economically-inactive, non-manual, manual), educational level (elementary school graduates and lower, junior high school graduates, high school graduates, college graduates and over), high-risk drinking (yes, no),

Income (quartile), smoking (current smoker, previous smoker, or non-smoker), skipped yesterday's breakfast (yes or no), skipped yesterday's lunch (yes or no), skipped yesterday's dinner (yes or no), dietary supplement consumption in the past one year (yes or no), usual fluid intake (g), protein intake (g), fat intake (g), carbohydrate intake (g), calculus intake (g), sodium intake (g), sinusitis prevalence (yes or no), otitis media prevalence (yes or no), tinnitus prevalence (yes or no), depression for two consecutive weeks (yes or no), pain and discomfort in the last two weeks (yes or no), and subjective voice problem recognition (yes or no).

C. Development and Validation of Prediction Models

This study developed models for predicting benign laryngeal disorders using deep learning, naive Bayes model, generalized linear model, CART, and random forest and compared the accuracy and runtime of them to check their prediction performance. Since this study had a small sample size ($n=626$), it could deteriorate the reliability when evaluating the prediction performance using held-out validation. Therefore, this study carried out 5-fold cross-validation to evaluate the prediction performance (Fig. 3). The R code of the 5-fold cross-validation is shown in Fig. 4.

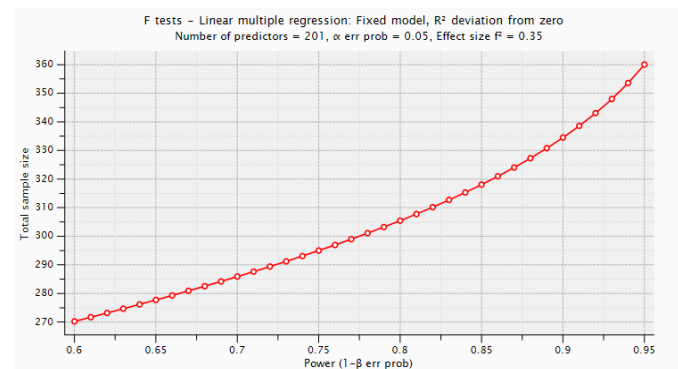


Fig. 1. Results of Power Test to Verify Statistical Significance Level.

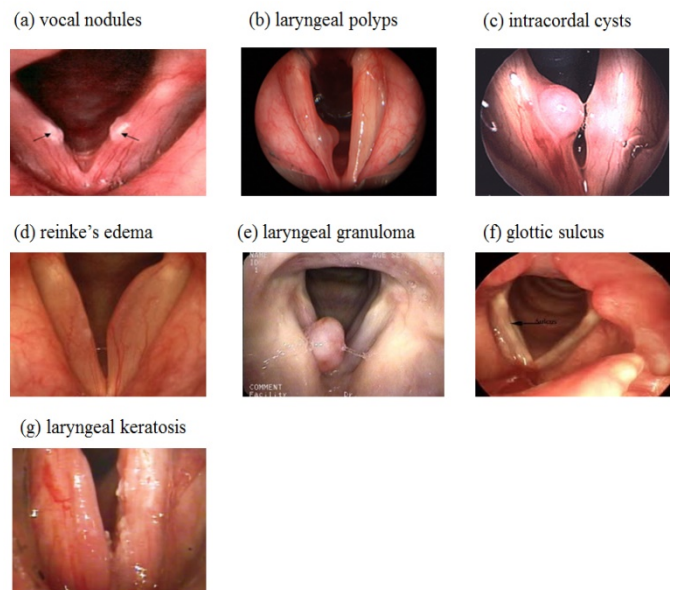


Fig. 2. Type of benign Vocal Fold Mucosal Disorders [38].

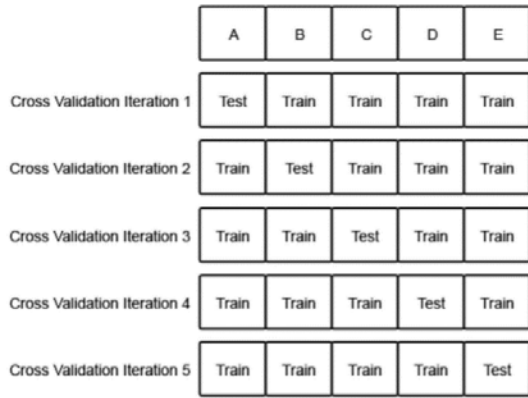


Fig. 3. The Concept of 5-fold Cross-validation.

```

glimpse(data)

#cross validation, using rf to predict sepal.length
k = 5

data$id <- sample(1:k, nrow(data), replace = TRUE)
list <- 1:k

# prediction and test set data frames that we add to with each iteration over
# the folds
prediction <- data.frame()
testsetCopy <- data.frame()

#Creating a progress bar to know the status of CV
progress.bar <- create_progress_bar("text")
progress.bar$init(k)
}
    
```

Fig. 4. R Code for 5-fold Cross-validation.

When developing a model using a method with a random characteristic (e.g., random forest), the seed was fixed to #0123456. This study defined the model with the highest accuracy as the model with the best prediction performance. When the accuracy was identical, a model with a shorter runtime was selected as the model with the best prediction performance. All analyzes were performed using R version 3.6.3 (Foundation for Statistical Computing, Vienna, Austria).

D. Machine Learning Models

The decision tree is an algorithm that creates a learning model in the tree shape according to the features of the data and derives a final decision through repetition. Since the decision tree expresses the analysis process in a tree-shaped graph, the decision tree has the advantage of helping a researcher understand and explain the analysis process easily (Fig. 5). In this study, CART was used as a decision tree algorithm. In this study, the maximum tree depth was set to 10, the parent node was set to 50, and the child node was set to 30.

The naive Bayes model is a method of classifying observations by using Bayes theory (Fig. 6). Bayes theory refers to a way of deriving a posteriori probability for a certain observation by using a secured prior probability.

Random forest is a decision tree-based ensemble method that generates many random samples using a bootstrap (randomly extracting samples of the same size from a given

data with replacement) from a learning data, trains independent decision trees for each sample group, and synthesizes the results to create a final model (Fig. 7).

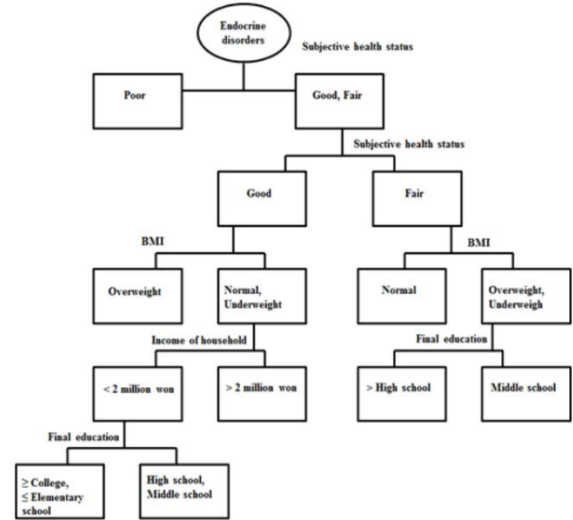


Fig. 5. Example of CART [39].

Input:

Training dataset T,
 $F = (f_1, f_2, f_3, \dots, f_n)$ // value of the predictor variable in testing dataset.

Output:

A class of testing dataset.

Step:

1. Read the training dataset T;
2. Calculate the mean and standard deviation of the predictor variables in each class;
3. Repeat

Calculate the probability of f_i using the gauss density equation in each class;

Until the probability of all predictor variables ($f_1, f_2, f_3, \dots, f_n$) has been calculated.

4. Calculate the likelihood for each class;
5. Get the greatest likelihood;

Fig. 6. Naive Bayes' Algorithm [40]

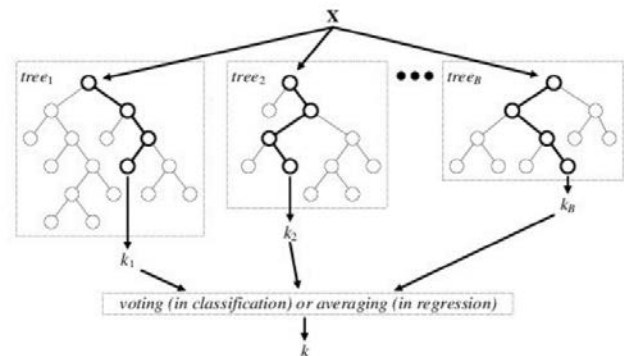


Fig. 7. Concept of Random Forest Model [41].

The generalized linear model is an extension of the linear model that can handle cases where a dependent variable of the dataset does not satisfy the normal distribution assumptions. It is a regression analysis using the glm() function. In other words, the generalized linear model models $f(x)$, which is formed by converting a dependent variable, using a linear combination of the independent variable and the regression coefficient.

Deep learning is an algorithm composed of an input layer, composed of independent variables, an output layer, composed of dependent variables, and two or more hidden layers between the input and output layers. Independent nodes are arranged in each layer, and the nodes between the two layers are connected by weighted neurons (connecting lines) (Fig. 8).

This study used H2O Deep Learning among various deep learning types. H2O's Deep Learning is a type of the multi-layer feedforward artificial neural networks, and it is trained with gradient descent optimization and back-propagation.

In this study, the number of hidden layers was set to 2 (200 hidden node), a default value, and epoch (the number of passes of the entire training dataset) was set to 10. H2O Deep Learning provides Tanh, Tanh with dropout, Rectifier, Rectifier with dropout, Maxout, and Maxout with dropout as activation functions. This study used rectifier, the default function, as an activation function to develop models. The code of H2O Deep Learning is presented in Fig. 9.

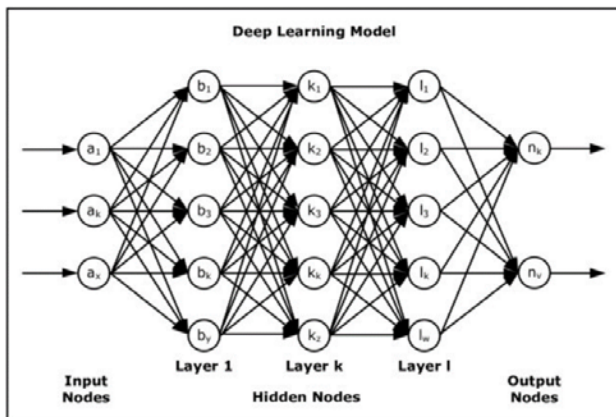


Fig. 8. The Concept of Deep Learning [42].

```
library(h2o)
h2o.init()

y = 'y_2002'
x = set_diff(names(train_data), y)

train.hex = as.h2o(train_data)
new_data.hex = as.h2o(test_data)

dl = h2o.deeplearning(x = x, y = y, training_frame = train.hex, nfolds = 5)

drf = h2o.randomForest(x = x, y = y, training_frame = train.hex, nfolds = 5)

glm = h2o.glm(x = x, y = y, training_frame = train.hex, nfolds = 5)
```

Fig. 9. Code of H2o Deep Learning.

III. RESULTS

A. Comparing the Accuracy and Runtime of benign Laryngeal Mucosal Disorder Prediction Models

The accuracies of five models (deep learning, naive Bayes model, generalized linear model, CART, and random forest) for predicting benign laryngeal mucosal disorders are presented in Fig. 10. In this study, deep learning was the best model with the highest accuracy (0.84). The runtimes of the five models are presented in Fig. 11. In this study, CART showed the shortest runtime (3min 7sec).

B. Predictors of benign Laryngeal Mucosal Disorders in Korean Adults

The normalized importance of the deep learning's variables, the final model, is presented in Fig. 12. This model confirmed that subjective voice problem recognition, pain and discomfort in the last two weeks, education level, occupation, mean monthly household income, high-risk drinker, and current smoker were major variables with high weight for the benign laryngeal mucosal disorders of Korean adults. Among them, subjective voice problem recognition was the most important factor with the highest weight.

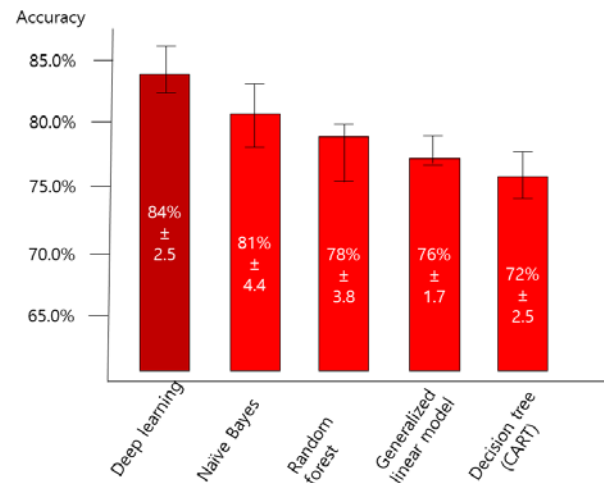


Fig. 10. Accuracy Comparison of Machine Learning and Deep Learning Models for Predicting benign Laryngeal Mucosal Disorders, (%).

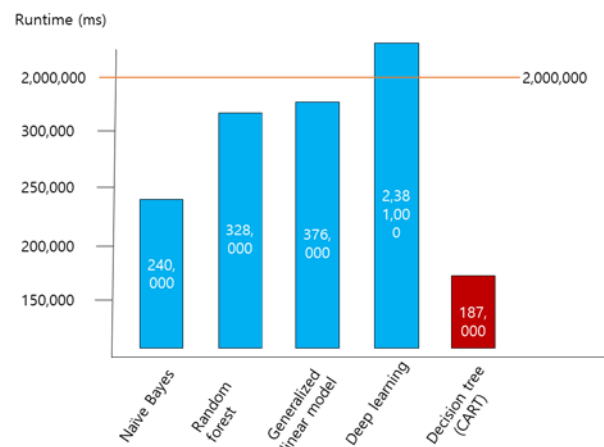


Fig. 11. Runtime Comparison of Machine Learning and Deep Learning Models for Predicting benign Laryngeal Mucosal Disorders, (ms).

Attribute	weight
Self-reported voice problem	0.094
Pain & discomfort over the last two weeks	0.075
Education level	0.066
Occupation	0.043
Average monthly household income	0.038

Fig. 12. Variable's Importance for benign Laryngeal Mucosal Disorders (Only Top 5).

IV. DISCUSSION

This study compared models for predicting the benign laryngeal mucosal disorders of Korean adults. The results of this study showed that deep learning had the best prediction performance among deep learning, naive Bayes model, generalized linear model, CART, and random forest. The runtime of deep learning was 39min 41sec, which was a 10 times longer development time than CART (3min 7sec). However, deep learning showed better ($\geq 6\%$) accuracy than machine learning models. The results of this study agreed with the results of previous studies [36, 43] that reported that the performance of deep learning was better than ensemble-based machine learning methods (e.g., light gradient boosted machine, and extreme gradient boosting) for predicting laryngeal disorders by using video, image, and speech analysis. The results of this study implied that the prediction performance of deep learning could be better than that of machine learning for structured data such as health behavior and demographic factors as well as video and image data. However, since machine learning studies using epidemiological data are much less than machine learning studies using video, image, and speech data, additional studies are needed to prove the superiority of prediction performance of deep learning in epidemiologic data such as health surveys.

V. CONCLUSION

The results of this study suggested that the prediction performance of deep learning could be better than other machine learning methods when developing a multi-modal model for predicting benign laryngeal mucosal disorders by using various data such as image data, demographic factors, and health behavior in the future. It will be necessary to compare the accuracy and runtime of models using the data of various diseases in order to prove the prediction performance of deep learning models, built by using epidemiological data.

ACKNOWLEDGMENT

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2018R1D1A1B07041091, NRF-2019S1A5A8034211).

REFERENCES

- [1] T. M. McCarrel, and J. B. Woodie, Update on laryngeal disorders and treatment. *Veterinary Clinics: Equine Practice*, vol. 31, no. 1, pp. 13-26, 2015.
- [2] H. S. Kim, Benign laryngeal disorders. *Korean Journal of Otorhinolaryngology-Head and Neck Surgery*, vol. 56, no. 6, pp. 332-338, 2013.
- [3] H. D. Soni, S. Gandhi, M. Goyal, and U. Shah, Study of clinical profile of benign laryngeal lesions. *International Journal of Medical Science and Public Health*, vol. 5, no. 4, pp. 656-660, 2016.

- [4] H. Byeon, and Y. Lee, Prevalence and risk factors of benign laryngeal lesions in the adult population. *Communication Sciences & Disorders*, vol. 15, no. 4, pp. 648-656, 2010.
- [5] M. M. Johns, Update on the etiology, diagnosis, and treatment of vocal fold nodules, polyps, and cysts. *Current Opinion in Otolaryngology & Head and Neck Surgery* vol. 11, no. 6, pp. 456-461, 2003.
- [6] H. Byeon, Model development for predicting the occurrence of benign laryngeal lesions using support vector machine: focusing on South Korean adults living in local communities. *International Journal of Advanced Computer Science and Applications*, vol. 9, no. 10, pp. 222-227, 2018.
- [7] N. Roy, R. M. Merrill, S. D. Gray, and E. M. Smith, Voice disorders in the general population: prevalence, risk factors, and occupational impact. *The Laryngoscope*, vol. 115, no. 11, pp. 1988-1995, 2005.
- [8] H. Byeon, and Y. Lee, Laryngeal pathologies in older Korean adults and their association with smoking and alcohol consumption. *The Laryngoscope*, vol. 123, no. 2, pp. 429-433, 2013.
- [9] H. Byeon, Gender differences in risk factors of benign vocal fold disease in Korea: the fifth Korea National Health and Nutritional Examination Survey. *Logopedics Phoniatrics Vocology*, vol. 41, no. 2, pp. 85-91, 2016.
- [10] H. Byeon, A population-based cross-sectional study of alcohol consumption and risk of benign laryngeal disease in Korean adults. *Journal of Voice*, vol. 30, no. 4, pp. 443-447, 2016.
- [11] S. H. Woo, R. B. Kim, S. H. Choi, S. W. Lee, and S. J. Won, Prevalence of laryngeal disease in South Korea: data from the Korea National Health and Nutrition Examination Survey from 2008 to 2011. *Yonsei Medical Journal*, vol. 55, no 2, p. 499-507, 2014.
- [12] H. Byeon, Occupational risks for voice disorders: evidence from a Korea national cross-sectional survey. *Logopedics Phoniatrics Vocology*, vol. 42, no. 1, pp. 39-43, 2017.
- [13] S. Simberg, A. Laine, E. Sala, and A. M. Rönnekaa, Prevalence of voice disorders among future teachers. *Journal of Voice*, vol. 14, no. 2, pp. 231-235, 2000.
- [14] M. Behlau, F. Zambon, A. C. Guerrieri, and N. Roy, Epidemiology of voice disorders in teachers and nonteachers in Brazil: prevalence and adverse effects. *Journal of Voice*, vol. 26, no. 5, pp. 665.e9-665.e18, 2012.
- [15] L. C. C. Cutiva, I. Vogel, and A. Burdorf, Voice disorders in teachers and their associations with work-related factors: a systematic review. *Journal of Communication Disorders*, vol. 46, no. 2, pp. 143-155, 2013.
- [16] H. Byeon, The prediction model for self-reported voice problem using a decision tree model. *Journal of the Korea Academia-Industrial cooperation Society*, vol. 14, no. 7, pp. 3368-3373, 2013.
- [17] 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.
- [18] 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.
- [19] 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 Otolaryngology, Rhinology & Laryngology*, vol. 124, no. 10, pp. 770-776, 2015.
- [20] R. M. B. De Alvear, F. J. Barón, and A. G. Martínez-Arquero, School teachers' vocal use, risk factors, and voice disorder prevalence: guidelines to detect teachers with current voice problems. *Folia Phoniatrica et Logopaedica*, vol. 63, no. 4, pp. 209-215, 2011.
- [21] 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.
- [22] H. Byeon, Comparative analysis of unweighted sample design and complex sample design related to the exploration of potential risk factors of dysphonia. *Journal of the Korea Academia-Industrial cooperation Society*, vol. 13, no. 5, pp. 2251-2258, 2012.
- [23] H. Byeon, Comparing the accuracy and developed models for predicting the confrontation naming of the elderly in South Korea using weighted

- random forest, random forest, and support vector regression. *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 2, pp. 326-331, 2021.
- [24] K. E. Bainbridge, N. Roy, K. G. Losonczy, H. J. Hoffman, and S. M. Cohen, Voice disorders and associated risk markers among young adults in the United States. *The Laryngoscope*, vol. 127, no.9, pp. 2093-2099, 2017.
- [25] L. M. da Rocha, S. de Lima Bach, P. L. do Amaral, M. Behlau, and L. D de Mattos Souza, Risk factors for the incidence of perceived voice disorders in elementary and middle school teachers. *Journal of Voice*, vol. 31, no. 2, pp. 258.e7-258.e12, 2017.
- [26] K. Kyriakou, K. Petinou, and I. Phiniketos, Risk factors for voice disorders in university professors in Cyprus. *Journal of Voice*, vol. 32, no. 5, pp. 643.e1-643.e9, 2018.
- [27] Y. R. Lee, H. R. Kim, and S. Lee, Effect of teacher's working conditions on voice disorder in Korea: a nationwide survey. *Annals of Occupational and Environmental Medicine*, vol. 30, no. 1, pp. 1-10, 2018.
- [28] H. Byeon, D. Lee, and S. Cho, Association between second-hand smoking and laryngopathy in the general population of South Korea. *PLOS ONE*, vol. 11, no. 11, pp. e0165337, 2016.
- [29] H. Byeon, D. Lee, and S. Cho, Relationship between women's smoking and laryngeal disorders based on the urine cotinine test: results of a national population-based survey. *BMJ Open*, vol. 6, no. 11, pp. e012169, 2016.
- [30] H. Byeon, A laryngeal disorders prediction model based on cluster analysis and regression analysis. *Medicine*, vol. 98, no. 31. pp. e16686, 2019.
- [31] Y. Wang, D. Wang, X. Ye, Y. Wang, Y. Yin, and Y. Jin, A tree ensemble-based two-stage model for advanced-stage colorectal cancer survival prediction. *Information Sciences*, vol. 474, pp. 106-124, 2019.
- [32] H. Byeon, Evaluating the accuracy of models for predicting the speech acceptability for children with cochlear implantations. *International Journal of Advanced Computer Science and Applications*, vol. 12, pp. 25-29, 2021.
- [33] H. Byeon, Comparing ensemble-based machine learning classifiers developed for distinguishing hypokinetic dysarthria from presbyphonia. *Applied Sciences*, vol. 11, no. 5, pp. 2235, 2021.
- [34] A. Korotcov, V. Tkachenko, D. P. Russo, and S. Ekins, Comparison of deep learning with multiple machine learning methods and metrics using diverse drug discovery data sets. *Molecular Pharmaceutics*, vol. 14, no. 12, pp. 4462-4475, 2017.
- [35] S. Bouktif, A. Fiaz, A. Ouni, and M. A. Serhani, Optimal deep learning lstm model for electric load forecasting using feature selection and genetic algorithm: comparison with machine learning approaches. *Energies*, vol. 11, no. 7, 1636, 2018.
- [36] H. Xiong, P. Lin, J. G. Yu, J. Ye, L. Xiao, Y. Tao, Z. Jing, W. Lin, M. Liu, J. Xu, W. Hu, Y. Lu, H. Liu, Y. Li, Y. Zheng, and H. Yang, Computer-aided diagnosis of laryngeal cancer via deep learning based on laryngoscopic images. *EBioMedicine*, vol. 48, pp. 92-99.
- [37] J. Ren, X. Jing, J. Wang, X. Ren, Y. Xu, Q. Yang, L. Ma, Y. Sun, W. Xu, N. Yang, J. Zou, Y. Zheng, M. Chen, W. Gan, T. Xiang, J. An, R. Liu, C. Lv, K. Lin, X. Zheng, F. Lu, Y. Rao, H. Yang, K. Liu, G. Liu, T. Lu, X. Zheng, and Y. Zhao, Automatic recognition of laryngoscopic images using a deep - learning technique. *The Laryngoscope*, vol. 130, no. 11, pp. E686-E693, 2020.
- [38] H. Byeon, Exploring potential risk factors for benign vocal fold mucosal disorders using weighted logistic regression. *International Journal of Bio-Science and Bio-Technology*, vol .6, no. 4, pp. 77-86, 2014.
- [39] 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.
- [40] M. F. A. Saputra, T. Widiyaningtyas, and A. P. Wibawa, Illiteracy classification using K means-Naïve Bayes algorithm. *JOIV: International Journal on Informatics Visualization*, vol. 2, no. 3, 153-158, 2018.
- [41] A. Verikas, E. Vaiciukynas, A. Gelzinis, J. Parker, and M. C. Olsson, Electromyographic patterns during golf swing: activation sequence profiling and prediction of shot effectiveness. *Sensors*, vol. 16, no. 4, pp. 592, 2016.
- [42] W. Serrano, Smart internet search with random neural networks. *European Review*, vol. 25, no. 2, pp. 260-272, 2017.
- [43] H. Kim, J. Jeon, Y. J. Han, Y. Joo, J. Lee, S. Lee, and S. Im, Convolutional neural network classifies pathological voice change in laryngeal cancer with high accuracy. *Journal of Clinical Medicine*, vol. 9, no. 11, pp. 3415, 2020.