Comparing the Accuracy of Deep Learning and Machine Learning for Predicting the Benign Laryngeal Disorders of Korean Adults

Haewon Byeon*, Shin Bong Gi**
*Department of Medical Big Data, College of AI Convergence, Inje University
**Department of Artificial intelligence, College of AI Convergence, Inje University
e-mail:byeon@inje.ac.kr

우리나라 성인의 후두양성점막질환 예측을 위한 딥러닝과 기계학습의 정확도 비교

변해원*, 신봉기**
*인제대학교 AI융합대학 메디컬 빅데이터전공
**인제대학교 AI융합대학 컴퓨터 공학부 인공지능전공

Abstract

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, 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). 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.

1. Introduction

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 [1,2,3]. Therefore, it is important in otolaryngology to accurately understand the etiology of a laryngeal disorder, diagnose it early, and provide appropriate treatment accordingly.

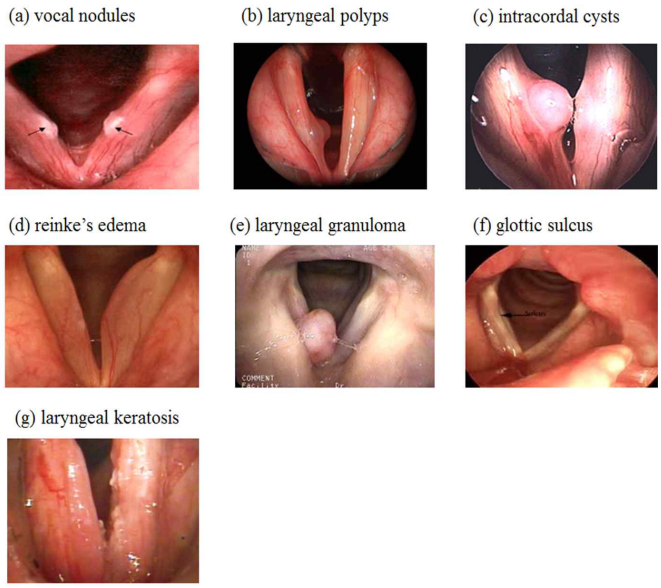
it is important to fully understand the etiology of a benign laryngeal mucosal disorder and identify multiple risk factors of the disease in order to perform accurate diagnosis and treatment. Nevertheless, most studies that have evaluated the risk factors of laryngeal disorders have just tried to find individual risk factors using regression analysis [4,5,6], and only a few studies have explored the multiple risk factors of benign laryngeal mucosal disorders using machine learning [7].

Supervised learning-based machine learning has been used as a way to detect a disease and identify multiple risks in recent years. Many recent studies [8,9] have reported that neural network-based deep learning is more accurate in classifying and predicting diseases than machine learning. Nevertheless, previous studies [10] 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 our study were to develop models for predicting benign laryngeal mucosal disorders based on deep learning, naive Bayes model, generalized linear model, 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.

2. Methods

This study targeted adults (≥19 years old) who participated in the otolaryngology examination and completed the 2012 KNHANES. 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). Benign laryngeal disease [11] in this study were defined as vocal nodules, laryngeal polyps, intracordal cysts, reinke's edema, laryngeal granuloma, glottic sulcus and laryngeal keratosis.



[Fig. 1] Type of benign vocal fold mucosal disorders [11]

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.

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 4.0.1 (Foundation for Statistical Computing, Vienna, Austria).

3. Results

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) in five models (deep learning, naive Bayes model, generalized linear model, CART, and random forest). The runtimes of the five models are presented in Fig. 11. In this study, CART showed the shorted runtime (3min 7sec). The normalized importance of the deep learning's variables, the final model, is presented in Fig. 2. 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.

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. 2] Variable's importance for benign laryngeal disorders (only top 5)

4. Conclusions

The results of our 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.

References

[1] S. Simberg, A. Laine, E. Sala, and A. M. Rönnemaa, "Prevalence of voice disorders among future teachers", Journal of Voice, vol. 14, no. 2, pp. 231-235, 2000.

[2] 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.

[3] 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.

[4] 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.

[5] K. Kyriakou, K. Petinou, and I. Phinikettos, "Risk factors for voice disorders in university professors in Cyprus", Journal of Voice, vol. 32, no. 5, pp. 643.e1-643.e9, 2018.

[6] 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.

[7] H. Byeon, "A laryngeal disorders prediction model based on cluster analysis and regression analysis", Medicine, vol. 98, no. 31. pp. e16686, 2019.

[8] 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.

[9] 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.

[10] 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.

[11] 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.