

POSTER PRESENTATION 4: INFORMATION AND COMMUNICATION

Development of the best ensemble-based machine learning classifier for distinguishing hypokinetic dysarthria caused by Parkinson's disease from presbyphonia and comparison of performance measures

H. Byeon

Purpose When people get old, they experience vocal aging due to the malfunction of the respiratory system and the vocal system. It is defined as presbyphonia in otorhinolaryngology. Presbyphonia generally shows symptoms such as hoarse, weak, or trembling voice due to the atrophy or loss of elasticity of the vocal cord muscles in the aging process. These symptoms are similar to the major vocal symptoms of early Parkinson's disease that are caused by damage to the nervous system. However, presbyphonia can be distinguished from neurological voice disorders such as vocal cord paralysis because presbyphonia is not a voice disorder. Therefore, it is essential to understand the aging process of voice characteristics to accurately distinguish presbyphonia from neurological voice disorders. This study developed the best ensemble-based machine learning classifier that could distinguish hypokinetic dysarthria from presbyphonia using classification and regression tree, random forest, GBM, and XGBoost, and compared the prediction performance of models. **Method** Voice was recorded using a microphone (D7 Vocal., AKG, Vienna, Austria) fixed at a 90-degree angle and installed at 10cm away from the mouth. Sustained vowels and connected speech were recorded at a sampling rate of 44,100Hz using the Analysis of Dysphonia in Speech and Voice (ADSV; Model 5109, Kay Pentax Medical, Montvale, NJ, USA). Sustained vowel phonation tasks were conducted by analyzing the phonation of /ah/ vowel in Korean for more than three seconds and the mean of three measurements was used for the analysis. Connected speech was collected by using "Gaul", a standardized paragraph test. The outcome variable was the prevalence of hypokinetic dysarthria. Explanatory variables were age, gender, years of education, current smoking status, current drinking status, CPP (dB), low-high spectral ratio(L/H ratio, dB), L/H ratio_standard deviation (SD), L/H ratio_max (dB), L/H ratio_Min (dB), CPP fundamental frequency (CPP F0, Hz), CSID, CPP Max (dB), CPP min (dB), mean CPP F0 (Hz), and mean CPP F0 SD (Hz). **Results and Discussion** This study compared the prediction performance of GBM, XGBoost, random forest, and classification and regression tree and found that the accuracy of random forests was the highest. It is believed that random forest showed higher accuracy than classification and regression tree because the former was based on a bagging algorithm that generated various decision trees from 500 bootstrap samples. However, the results of this study revealed that the sensitivity of XGBoost was higher than that of random forest. Therefore, future studies are necessary to compare prediction performance by estimating diverse evaluation methods, such as sensitivity, specificity, and weight harmonic average, suitable for the analysis objective, rather than using one performance index such as accuracy. The results of this study imply that CPP among Cepstral indices is useful for distinguishing hypokinetic dysarthria from presbyphonia in acoustic-phonetic analysis using both Cepstrum and spectrum. It will be necessary to develop a multimodal-based prediction model including auditory-perceptual indices, biomarkers, and acoustic-phonetic indices (e.g., Ceptral and spectral measures) to predict hypokinetic dysarthria more sensitively.

Keywords: presbyphonia, Parkinson's disease, Hypokinetic dysarthria, ceptral, spectral

Address: Inje University, Republic of Korea

Email: bhwpuma@naver.com

Acknowledgement: This research was funded by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, grant number "2018R1D1A1B07041091, 2021S1A5A8062526", and "2022 Development of Open-Lab based on 4P in the Southeast Zone".