Evaluating the Accuracy of Models for Predicting the Speech Acceptability for Children with Cochlear Implants

Haewon Byeon

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

Abstract—This study developed a model for predicting healthy hearing people's speech acceptability for children with cochlear implants using multiple regression analysis, support vector regression, and random forest and evaluated the prediction performance of the model by comparing mean absolute errors and root mean squared errors. This study targeted 91 hearing-impaired children between four and eight years old who had worn cochlear implants at least one year and less than five years. Speech data of children wearing cochlear implants (CI) were collected through two tasks: speaking and reading. The outcome variable, healthy hearing people's speech acceptability for children wearing CI was evaluated by 80 college students (freshman and sophomore) who did not have prior knowledge of children with a cochlear implant. The results of this study showed that the random forest algorithm (mean absolute errors=0.81and root mean squared error=0.108) was the best model for predicting the speech acceptability of children wearing CI. The results of this study imply that the predictive performance of random forest will be the best among ensemble models when developing a machine learning model using speech data of children wearing CI.

Keywords—Cochlear implants; speech acceptability; support vector regression; random forest; mean absolute errors

I. INTRODUCTION

Since the National Health Service of South Korea began to cover cochlear implants in 2005, cochlear implants have become more common for the hearing impaired in South Korea. The ear consists of the external ear, the middle ear, and the internal ear, and the cochlear implantation refers to an operation of implanting an artificial cochlea device in the ear of the patient who cannot hear voice due to the damage of the cochlea to help the patient hear speech [1]. Cochlear implants provide useful hearing for children with severe hearing difficulties or deaf children who cannot hear speech even with hearing aids [2, 3]. Many children with hearing impairments have benefited greatly from cochlear implants (CI) for enhancing their hearing ability and developing language ability [4]. Particularly, the ultimate goal to obtain through cochlear implants is to improve communication skills through vocal language [5]. Consequently, many studies [6, 7] have shown interest in the ability of children to produce spoken language (speech) after cochlear implants, and they have sought ways for children with cochlear implants to produce better speech than before operation based on the improved hearing ability.

Speech intelligibility and speech acceptability have been used widely in the speech-language pathology field to compare speech characteristics and severity for diverse communicative disorders such as articulation and phonological disorders, dysarthria, and apraxia of speech [8, 9]. Among them, speech acceptability refers to how well the content that the speaker is trying to convey is delivered to the listener (how well the listener understands it", and it is mainly used as an index reflecting the success of expressive speech [8]. Since the listener listens to the speaker, various variables (e.g., vocal intensity, pitch, and speech rate) comprehensively influencing the listening. It is necessary to have an index for evaluating the overall speech production ability of the speaker from the listener's point of view [5]. Speech acceptability is used as an index to evaluate the overall speech production ability. Speech acceptability measures how naturally the speaker's intention is understood by the listener, and is a representative index showing the expressive ability of the hearing impaired.

Nevertheless, previous studies measured speech acceptability to mainly evaluate the speech characteristics of patients with dysarthria or the hearing-impaired due to neurological damage such as stroke and to identify the speech characteristics of cleft palate patients [8, 9, 10]. These studies compared the results with the speech acceptability of the healthy control group based on traditional statistical analyses such as t-test and ANOVA [8, 9, 10]. Only a few studies examined the predictors of speech acceptability using machine learning.

The general public has become more interested and gained a better understanding in machine learning in various fields (e.g., finance, medicine, and engineering) [11, 12]. The South Korean government also pays a lot more policy interest in the artificial intelligence (AI) and health care industries. AI refers to a technology for analyzing data and identifying better answers through the visualization of big data, machine learning, and deep learning of big data. Among them, machine learning indicates a prediction technique to predict changes by reading numerous data and discovering hidden algorithms. In the healthcare industry, there have been many cases of applying and utilizing AI technologies including machine learning [13], such as cancer diagnosis and treatment recommendations using AI-based IBM Watson, diagnostic medicine using machine learning analysis techniques, and new drug development systems. Studies have continuously reported

that models relying on machine learning had better prediction power than traditional statistical techniques based on general linear model (GLM) [14, 15, 16, 17]. It is still necessary to develop machine learning-based prediction models and compare their prediction power with the prediction power of GLM-based regression models for proving the usefulness of machine learning in the medical field. This study developed a model for predicting healthy hearing people's speech acceptability for children with cochlear implants using multiple regression analysis, support vector regression, and random forest and evaluated the prediction performance of the model by comparing mean absolute errors and root mean squared errors.

II. RESEARCH METHODS

A. Study Subjects

It is a descriptive study that identified the factors associated with the speech acceptability for children with cochlear implants perceived by people with healthy hearing and this study targeted 91 hearing-impaired children between four and eight years old who resided in Seoul, Incheon, and Suwon and had worn cochlear implants at least one year and less than five years. The subjects of this study were the same as Byeon [5]. The study subjects were (1) the hearing-impaired who had worn cochlear implants at least one year, (2) those who received hearing rehabilitation regularly after surgery, and (3) those who were using oral speech during a conversation. This study excluded children with a cognitive disorder, an affection disorder, visual impairment, Autism spectrum, and development disabilities in addition to hearing impairment. The power was tested using G-Power version 3.1.9.7 (Universität Mannheim, Mannheim, Germany) (Fig. 1), and the minimum sample size was derived as 80 samples when the number of predictors was seven, the significance level was a=0.05, power (1-B) was 0.8, and the effect size (f2) was 0.2. This study's sample size satisfied the appropriate sample size for testing the statistical significance (Fig. 2).
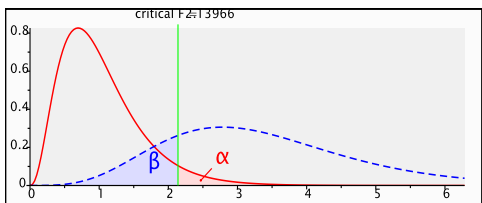

Fig. 1.   Result of Power Analysis.

Fig. 2.   Results of Sample Size Calculation.

B. Measurement

Speech data of children wearing cochlear implants (CI) were collected through two tasks, speaking and reading. The reading sentence was "Once upon a time, there was a young tiger living in a village. The young tiger was very curious" referring to Yoon [18]. The speaking task was to introduce oneself in the form of "Nice to meet you. My name is OOO." It was recorded using the Multi-Dimensional Voice Program (MDVP, Key Pentax, USA) installed on the computer in a quiet room without noise, and the microphone (Shure BETA58A) used for recording was located 10cm below the child's mouth.

The outcome variable, healthy hearing people's speech acceptability for children wearing CI, was evaluated by 80 college students (freshman and sophomore) who did not have prior knowledge of children with a cochlear implant. Each evaluator evaluated the speech acceptability of each child after listening to the speech data of the child once, which was played on a computer through a speaker in a noise-free place, and there was a 5-second interval between speech data. Speech acceptability was measured using a visual analog scale. The evaluators indicated the degree of speech acceptability perceived by them on a 100mm straight line where 0 was marked as "impossible to understand" and 100 was marked as "fully understandable" [19]. After the first evaluation was completed, the second evaluation was performed by changing the presentation order of the speech data. The mean values of the first and second evaluations were defined as the final scores of speech acceptability for the subjects' reading and speaking.

Explanatory variables included gender, age, household income, the period of wearing cochlear implants, corrected hearing, auditory-language rehabilitation period, pitch, loudness, and quality. Corrected hearing was defined as the mean threshold decibels (dB) of hearing tests measured in the ranges of 250, 500, 1k, 2k, and 4kHz after wearing a cochlear implant. Where the subject wore cochlear implants for both ears, the mean threshold value was used. When a cochlear implant was used for one ear and a hearing aid was used for the other ear, only the hearing of the cochlear implant side was used. Pitch, loudness, and quality were defined by analyzing the speech data recorded in MDVP.

C. Analysis Methods

This study developed a model for predicting the speech acceptability of children wearing CI using multiple regression analysis, support vector regression analysis, and random forest algorithm. This study also evaluated and validated the model to test the prediction performance of the developed model. This study randomly divided the data into a training dataset (70%) and a test dataset (30%) for validating the prediction performance; the training dataset was used to develop a prediction model and the test dataset was used to evaluate the prediction performance (mean absolute error and root mean squared error) by using the test dataset. All analyses were performed using R version 4.0.2 (Foundation for Statistical Computing, Vienna, Austria) and Python version 3.8.0 (https://www.python.org). The schematic diagram of the study is presented in Fig. 3.
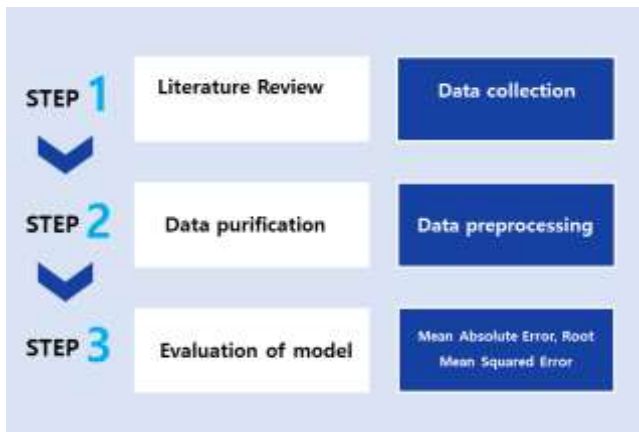
Fig. 3.   The Schematic Diagram of the Study.

D. Multiple Regression Analysis

Multiple regression analysis is an analysis that models the relationship between data while reiterating the process to minimize the error between the given data and the values obtained by the selected learning model. Linear regression is a method of analyzing the linear relationship between a dependent variable and at least one independent variable. When the dependent variable is a continuous variable, it can be analyzed using linear regression. When using the multiple linear regression analysis, it is possible to identify the influence (weight) of each independent variable on the dependent variable by estimating the regression coefficient. The least squares method or the maximum likelihood estimation method is used to estimate regression coefficients when modeling using the multiple linear regression method to predict results. Generally, the least squares method is used to make a regression model and analysis prediction results. The least squares method uses a method that minimizes the error of the model (the difference between estimated values of a model and actual observations) for estimate regression coefficients. Therefore, it searches for a model that can estimate values that are close to the actual results. This study also constructed a multiple linear regression model by applying the least squares method. An example of the least squares method is presented in Fig. 4.
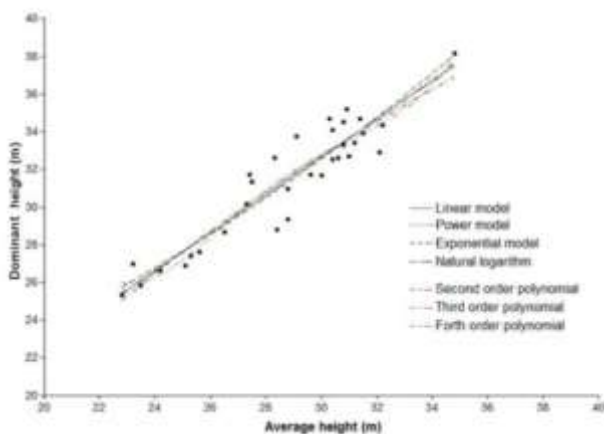

Fig. 4.   Examples of Regression Analysis by Least Squares Method with different Model Formulations [20].

E. Random Forest

Random forest is composed of multiple decision trees. The goal of random forest is to make more accurate predictions by making multiple decision tree models. Random forest is a decision tree-based ensemble method, which generates numerous random samples through a bootstrap method that randomly extracts samples with replacement of the same sample size from the training dataset, learns an independent decision tree for each sample dataset, and determines the final model by summarizing the results. The ensemble method is to create a final prediction model by generating multiple prediction models from a given data and then combining them. Many previous studies have shown that the ensemble method can improve the predictive power of the model [21,22]. Moreover, random forest has smaller prediction errors with more decision trees and it does not overfit even if there are many decisions, which are advantages of random forest. The concept of random forest is presented in Fig. 5.
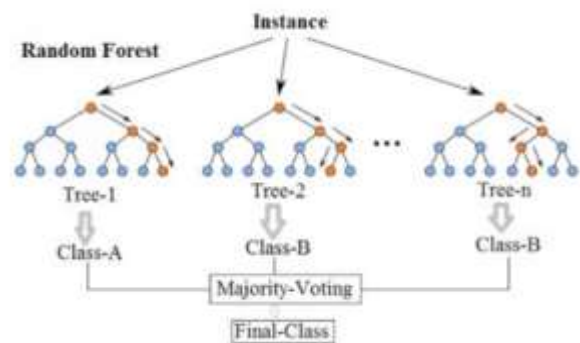

Fig. 5.   The Concept of Random Forest [23].

F. Support Vector Regression Analysis

Support vector regression is a regression model based on a support vector machine (SVM). SVM finds the optimal hyperplane that classifies the data into the most suitable classes by maximizing the margin for classifying input data by expressing the data in a high-dimensional vector space using a kernel function. Support vector regression is an extension of this SVM so that SMV can be applied to regression analysis. It is used to predict a random error tolerance value by introducing an e-insensitive loss function [24]. Support vector regression, like SVM, uses a kernel function to converting training data into points in feature space and then performs learning in feature likelihood. However, SVM and support vector regression are different in the aspect that SVM is a machine learning to classify "+1 class" and "-1 class", while support vector regression is a method to generalize class for predicting random error tolerance values using a regression function [25]. Support vector regression has the advantage of having high explanatory power even for data showing nonlinearity or complex patterns. However, it takes a long time to learn because computational complexity is high and it is not easy to interpret the model because it is not possible to analyze the direct relationship between independent and dependent variables, which are disadvantages. Moreover, support vector regression converts a nonlinear feature dimension, which cannot be linearly separated linearly, into a high-dimensional linear regression problem using a kernel function for nonlinear

expansion. The kernel function generally used in this process is a linear, polynomial, or radial basis function. The concept of support vector regression is presented in Fig. 6.
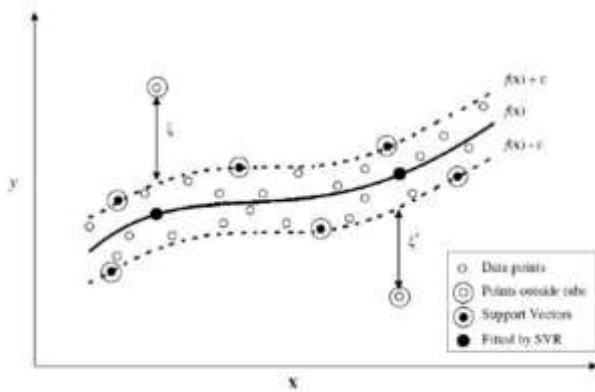

Fig. 6.   The Concept of Support Vector Regression [26].

G. Evaluating the Prediction Performance of the Model

A multiple linear regression model was built by using a regression coefficient estimation method based on the least squares method. While conducting random forest analysis, the limit of decision tree development was set to 100. Support vector regression was analyzed using a linear kernel function, the most basic kernel function, c (the parameter determining the generalization of a regression model) was 15.0, and the e-insensitive loss function (a precision parameter) was set as 0.001. This study compared mean absolute errors and root mean squared errors to evaluate the prediction performance of the developed models. Since random forest includes randomness, the model was developed while fixing the seed (#123456) during repeated measurements.

III.   RESULTS

A. Comparing the Performance of Models for Predicting Healthy hearing People's Speech Acceptability for Children Wearing CI

Table I shows the mean absolute errors and root mean squared errors of the speech acceptability prediction model for children wearing CI using multiple regression analysis, support vector regression analysis, and random forest. This study defined that a model with the smallest mean absolute error and root mean squared error was the best model with the best prediction performance. The results of this study showed that the random forest algorithm (mean absolute errors=0.81and root mean squared error=0.108) was the best model for predicting the speech acceptability of children wearing CI.

B. The Importance of Variables in the Final Model (Random Forest) for Predicting the Speech Acceptability for Children with CI Wearers

The normalized importance of random forest variables (the final model) is presented in Fig. 7. It was found that pitch, loudness, quality, the duration of wearing cochlear implants, the duration of aural rehabilitation, corrected hearing, and age were major variables with high weight in predicting the speech acceptability of children wearing CI. Among these variables, pitch was the most important factor in the final model.

TABLE I.   MEAN ABSOLUTE ERRORS AND ROOT MEAN SQUARED ERRORS OF MODELS FOR PREDICTING THE SPEECH ACCEPTABILITY FOR CHILDREN WEARING CI USING MULTIPLE REGRESSION ANALYSIS, SUPPORT VECTOR REGRESSION, AND RANDOM FOREST

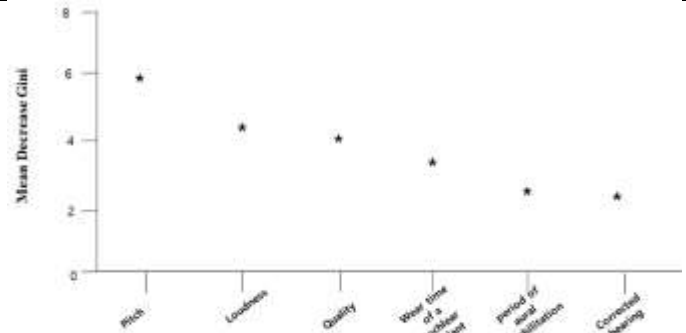| Type of model	Mean absolute error	Root mean squared error
Multiple regression	0.084	0.113
Support vector regression	0.081	0.109
Random forest	0.081	0.108

Fig. 7.   The Normalized Importance of Variables in the Model for Predicting Healthy Hearing People's Speech Acceptability for Children Wearing CI based on Random Forest (Only the Results of the Top 6 Variables are Presented).

IV.   CONCLUSION

This study developed a model for predicting healthy hearing people's speech acceptability for children wearing CI and found that pitch, loudness, and quality were main variables with higher weight for predicting the speech acceptability of children wearing CI. Among them, pitch was the most important factor in the final model. Factors affecting speech acceptability can be divided into segmental factors such as the errors in individual consonants and vowels and supra-segmental factors such as stress, speaking rate, voice quality, and intensity. Dagenais et al. [27] evaluated dysarthria and showed that speech acceptability was significantly correlated with speaking rate. Moreover, Lee et al. [19] analyzed the speech acceptability of hearing-impaired adults and reported that the speech acceptability of them was more strongly correlated with supra-segmental factors than segmental factors, and consonant accuracy, intonation, resonance, and speech rate were major variables influencing speech acceptability. Previous studies [28, 29] that analyzed the acoustic and phonetic characteristics of speech made by hearing impaired children with wearing CI showed that the pitch and quality related indices of children wearing CI were different from those of healthy hearing children. Hsu et al. [30], who evaluated auditory senses, also showed that the speech characteristics of children wearing CI were different from those of healthy hearing children in terms of pitch, quality, and resonance. In summary, the results of this study suggested that the speech characteristics of hearing impaired children with wearing CI, which the listener felt unnatural, were mostly due to acoustic-phonetic characteristics such as pitch and loudness among various speech-related factors such as age and gender.

Another finding of this study was random forest had the best prediction performance among multiple regression

analysis, support vector regression analysis, and random forest after comparing the accuracy of the models for predicting the healthy hearing people's speech acceptability for children wearing CI. This study developed prediction models using random forest (a machine learning technique), support vector regression analysis (a machine learning technique), and multiple regression analysis (a GLM analysis technique) and evaluated prediction performance by calculating mean absolute errors and root mean squared errors. The results of this study showed that random forest-based speech acceptability prediction model for children wearing CI showed the smallest mean absolute error and root mean squared error among the three models. This result agreed with the results of previous studies [14, 15, 16, 17] indicating that random forest-based models performed better than regression models in predicting diseases. The results of this study support the possibility that the accuracy of the ensemble model may be better than that of GLM. Furthermore, they imply that the predictive performance of random forest will be the best among ensemble models when developing a machine learning model using speech data of children wearing CI. Further studies are needed to prove the prediction performance of random forest by comparing accuracy using data from various fields.

REFERENCES

[1] M. Castro-Neto, Y. Jeong, M. K. Jeong, and L. D. Han, AADT prediction using support vector regression with data-dependent parameters. Expert Systems with Applications, vol. 36, no. 2, pp. 2979-2986, 2009.

[2] M. A. Svirsky, A. M. Robbins, K. I. Kirk, D. B. Pisoni, and R. T. Miyamoto, Language development in profoundly deaf children with cochlear implants. Psychological Science, vol. 11, no. 2, pp. 153-158, 2000.

[3] H. Byeon, Development trends of online-based aural rehabilitation programs for children with cochlear implant coping with the fourth industrial revolution and implication in speech-language pathology. International Journal of Advanced Computer Science and Applications, vol. 10, no. 10, pp. 25-31, 2019.

[4] H. Bruijnzeel, F. Ziylan, I. Stegeman, V. Topsakal, and W. Grolman, Systematic review to define the speech and language benefit of early. Audiology and Neurotology, vol. 21, no. 2, pp. 113-126, 2016.

[5] H. Byeon, Developing a model for predicting the speech intelligibility of South Korean children with cochlear implantation using a random forest algorithm. International Journal of Advanced Computer Science and Applications, vol. 9, no. 11, pp. 88-93, 2018.

[6] M. C. Allen, T. P. Nikolopoulos, and G. M. O'Donoghue, Speech intelligibility in children after cochlear implanation. Otology & Neurotology, vol. 19, no. 6, pp. 742-746, 1998.

[7] A. E. Geers, Speech, language, and reading skills after early cochlear implantation. Archives of Otolaryngology–Head & Neck Surgery, vol. 130, no. 5, pp. 634-638, 2004.

[8] T. L. Whitehill, Assessing intelligibility in speakers with cleft palate: a critical review of the literature. The Cleft Palate-Craniofacial Journal, vol. 39, no. 1, pp. 50-58, 2002.

[9] L. W. Ellis, Magnitude estimation scaling judgments of speech intelligibility and speech acceptability. Perceptual and Motor Skills, vol. 88, no. 2, pp. 625-630, 1999.

[10] I. K. Y. Law, E. P. M. Ma, and E. M. L. Yiu, Speech intelligibility, acceptability, and communication-related quality of life in Chinese alaryngeal speakers. Archives of Otolaryngology–Head & Neck Surgery, vol. 135, no. 7, pp. 704-711, 2009.

[11] M. Chen, Y. Hao, K. Hwang, L. Wang, and L. Wang, Disease prediction by machine learning over big data from healthcare communities. Ieee Access, vol. 5, pp. 8869-8879, 2017.

[12] H. Byeon, Best early-onset Parkinson dementia predictor using ensemble learning among Parkinson's symptoms, rapid eye movement sleep disorder, and neuropsychological profile. World Journal of Psychiatry, vol. 10, no. 11, pp. 245-259, 2020.

[13] Y. Lee, and J. Kim, Artificial intelligence technology trends and IBM Watson references in the medical field. Korean Medical Education Review, vol. 18, no. 2, pp. 51-57, 2016.

[14] H. Byeon, Development of a depression in Parkinson's disease prediction model using machine learning. World Journal of Psychiatry, vol. 10, no. 10, pp. 234-244, 2020.

[15] M. Khalilia, S. Chakraborty, and M. Popescu, Predicting disease risks from highly imbalanced data using random forest. BMC Medical Informatics and Decision Making, vol. 11, no. 1, pp. 51, 2011.

[16] H. Byeon, Exploring the predictors of rapid eye movement sleep behavior disorder for Parkinson's disease patients using classifier ensemble. Healthcare, vol. 8, no. 2, pp. 121, 2020.

[17] H. Byeon, Application of machine learning technique to distinguish Parkinson's disease dementia and Alzheimer's dementia: predictive power of Parkinson's disease-related non-motor symptoms and neuropsychological Profile. Journal of Personalized Medicine, vol. 10, no. 2, pp. 31, 2020.

[18] M. S. Yoon, Variables for predicting speech acceptability of children with cochlear implants. Phonetics and Speech Sciences, vol. 6, no. 4, pp. 171-179, 2014.

[19] S. E. Lee, H. H. Sim, C. M. Nam, J. Y. Choi, and E. S. Park, Auditory-perceptual evaluation of the speech of hearing-impaired adults: based on suprasegmental factors, speech intelligibility, and speech acceptability. Communication Sciences & Disorders, vol. 15, no. 4, pp. 477-493. 2010.

[20] A. Ferezliev, R. Mavrevski, and A. Delkov, Correlation between average and dominant height of middle-aged douglas fir plantations in the North-West Rhodopes. Silva Balc, vol. 19, no. 2, pp. 13-26, 2018.

[21] P. T. Noi, and M. Kappas, Comparison of random forest, k-nearest neighbor, and support vector machine classifiers for land cover classification using Sentinel-2 imagery. Sensors, no. 18, no. 1, pp. 18, 2018.

[22] H. Byeon, Is the random forest algorithm suitable for predicting Parkinson's disease with mild cognitive impairment out of Parkinson's disease with normal cognition?. International Journal of Environmental Research and Public Health, vol. 17, no. 7, pp. 2594, 2020.

[23] A. Chakure, Random Forest Classification. The Stratup, 2019.

[24] A. J. Smola, and B. Schölkopf, A tutorial on support vector regression. Statistics and computing, vol. 14, no. 3, pp. 199-222, 2004.

[25] M. Castro-Neto, Y. Jeong, M. K. Jeong, and L. D. Han, AADT prediction using support vector regression with data-dependent parameters. Expert Systems with Applications, vol. 36, no. 2, pp. 2979-2986, 2009.

[26] S. K. Lahiri, and K. C. Ghanta, The support vector regression with the parameter tuning assisted by a differential evolution technique: Study of the critical velocity of a slurry flow in a pipeline. Chemical Industry and Chemical Engineering Quarterly, vol. 14, no. 3, pp. 191-203, 2008.

[27] P. A. Dagenais, G. R. Brown, and R. E. Moore, Speech rate effects upon intelligibility and acceptability of dysarthric speech. Clinical Linguistics & Phonetics, vol. 20, no. 2/3, pp. 141-148, 2006.

[28] M. S. Yoon, E. A. Choi, and Y. J. Sung, A Comparison of voice analysis of children with cochlear implant and with normal hearing. Journal of the Korean Society of Speech Sciences, vol. 5, no. 4, pp. 71-78, 2013a.

[29] L. H. P. Nguyen, J. Allegro, A. Low, B. Papsin, and P. Campisi, Effect of cochlear implantation on nsality in children. Ear, Nose & Throat Journal, vol. 87, no. 3, pp. 138-143, 2008.

[30] H. W. Hsu, T. J. Fang, L. A. Lee, Y. T. Tsou, and S. H. Chen, Multidimensional evaluation of vocal quality in children with cochlear implants: a cross-sectional, case-controlled study. Clinical Otolaryngology, vol. 39, no. 1, pp 32-38, 2014.