

Automated Dysarthria Severity Classification Using Diadochokinetic test and Speech Intelligibility Based on LightGBM

MinSeo Kim¹, Junseok Oh² and Ji-Hwan Kim^{2*}

¹ Department of Artificial Intelligence, Sogang University Seoul, South Korea

² Department of Computer Science and Engineering, Sogang University Seoul, South Korea
[e-mail: minseokim@sogang.ac.kr, ohjs@sogang.ac.kr, kimjihwan@sogang.ac.kr]

*Corresponding author: Ji-Hwan Kim

Abstract

There has been a growing interest in developing digital treatments for dysarthria recently. Assessing the severity of dysarthria and identifying its cause are crucial for personalizing patient treatment. In this paper, we use NeuroSpeech source code to extract Diadochokinetic (DDK) features. We propose a severity classification model based on LightGBM, using the extracted DDK features, gender, and intelligibility. DDK regularity, duration, energy, and 8 other features were extracted from the diadochokinetic test, where /pa/, /ta/, and /ka/ were repeatedly pronounced. We trained a model to classify severity level with the gender, intelligibility and DDK features extracted from the speech of 320 patients, and it showed an accuracy of 79.17%. Intelligibility, number of pauses per second and average of the syllables duration have the high importance on the severity level.

Keywords: Diadochokinetic Test, Dysarthria, LightGBM, NeuroSpeech

1. Introduction

Dysarthria refers to a group of speech impairments caused by damage to the peripheral nervous system or central nervous system [1]. Some people who have had a stroke may suffer from dysarthria. These symptoms can often be alleviated through ongoing speech rehabilitation therapy. Research has been underway to develop digital treatments that can be customized to address individualized need or challenges. In order to provide personalized treatment, it is necessary to assess the severity of dysarthria and determine which assessment factors are problematic. A diadochokinetic (DDK) test can provide information about the condition of patients. We can use the features extracted by DDK test to assess severity and provide personalized care.

Speech intelligibility refers to the extent to which listeners understand what the speaker is trying to

communicate. Intelligibility offers information about the overall severity of dysarthria. We utilized intelligibility as a classification feature for assessing severity with DDK features.

In this paper, we extract severity diagnostic features from the diadochokinetic speech of stroke patients with dysarthria and propose a model that can classify severity level based on lightGBM.

2. Related Works

2.1 Diadochokinetic

A diadochokinetic test requires an individual to repeatedly and quickly pronounce /pa/, /ta/, and /ka/ within a specified timeframe. The diadochokinetic test is a fundamental component of motor speech assessment, providing quantitative measurements of the speech articulators' motor abilities [1].

This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government(MSIT) (No.2022-0-00621, Development of artificial intelligence technology that provides dialog-based multi-modal explainability)

2.2 NeuroSpeech

NeuroSpeech is an open source software that analyzes speech signals from dysarthria patients with Parkinson's disease [2]. The software can analyze speech signals in five dimensions : phonation, articulation, prosody, DDK, and intelligibility [2]. The features extracted from the DDK analysis are related with DDK regularity, duration, energy, pause rate, and 8 other features.

2.3 LightGBM

LightGBM is a powerful machine learning algorithm that can be utilized for various tasks, including feature selection and regression analysis [3]. It utilizes a unique combination of gradient-based one-side sampling (GOSS) and histogram-based gradient boosting techniques, allowing for faster training and prediction times while effectively handling high-dimensional features.

3. Experiments

The dataset¹ used in this study consists of 957 diadochokinetic speech recordings from 320 stroke patients with dysarthria. Severity level was measured on a scale of mild, moderate, and severe, and intelligibility was measured on a scale of 1-5.

We extracted DDK features using the NeuroSpeech source code and trained the lightGBM model using these features, along with gender, age, and intelligibility. We used 80% of the dataset as train data, 10% as validation data, and 10% as test data. Since the data is unbalanced (794 mild, 109 moderate, and 54 severe), we weighted each class ratio when training the LightGBM model.

We measured the model accuracy with test data and it shows 79.17% accuracy. We used SHAP value as the visualization method. **Fig. 1** shows which features are important for classifying severity level. Intelligibility, number of pauses per second (Pause_rate), and average of the syllables duration (avg_dur_ddk) have high importance. Intelligibility, not DDK features, has the highest importance. When evaluating the test data with the model trained using DDK features

except intelligibility, it showed 61.46% accuracy. This shows that intelligibility has the highest importance, but DDK features are also important for classifying severity.

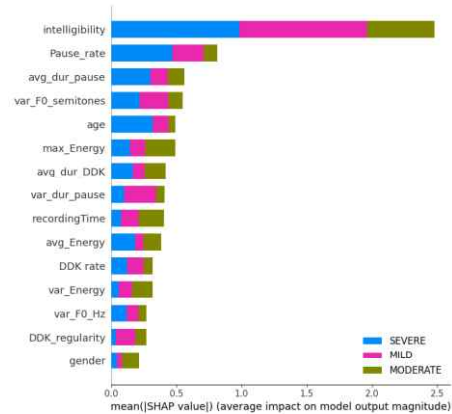


Fig. 1. Feature importance for severity classification

4. Conclusions

In this paper, we extracted features from diadochokinetic speech of stroke patients with dysarthria and use them to classify severity level with lightGBM model. We also investigated which features have high importance in classifying severity. The model demonstrated 75% accuracy, with intelligibility being the most important factor. However, the model performance improved when DDK features were included. Future work will focus on studying the automatic assessment of intelligibility using DDK speech, with the aim of classifying severity using only speech data.

References

- [1] F. L. Darley, A. E. Aronson, and J. R. Brown, "Differential diagnostic patterns of dysarthria," *Journal of Speech and Hearing Research*, vol. 12, no. 2, pp. 246-269, 1969.
- [2] J. R. Orozco-Arroyave et al., "NeuroSpeech: An open-source software for Parkinson's speech analysis," *Digital Signal Processing*, vol. 77, pp. 207-221, 2018.
- [3] G. Ke et al., "LightGBM: A highly efficient gradient boosting decision tree," *Advances in Neural Information Processing Systems*, vol. 30, 2017.

¹ This paper used datasets from 'The Open AI Dataset Project (AI-Hub, S. Korea)'. All data information can be accessed through 'AI-Hub (www.aihub.or.kr)'.