IOCBE 2024 Conference
The 1st International Online Conference on Bioengineering
16–18 October 2024 | Online
MDPI

Detection of Developmental Language Disorder Using Machine Learning and Mel-Frequency Cepstral Coefficients from Voice Recordings

Francesco Romano, Daniela Cardone, Arcangelo Merla, David Perpetuini
Department of Engineering and Geology, Unich, 66100, Italy

INTRODUCTION & AIM

Language disorder is the most frequent developmental disorder in childhood, impacting various aspects of language processing. Approximately 11-18% of children aged 18-36 months exhibit language delays, often improving by age 3, but some persist into developmental language disorder (DLD). The early detection of DLD is crucial, as it allows for timely intervention, improving long-term outcomes. This study aims to assess DLD though machine learning (ML) techniques applied to Mel-Frequency Cepstral Coefficients (MFCCs), which are features commonly used in voice analysis [1].

METHOD

This study utilized a freely available dataset [2] that comprises voice recordings from 54 children (35 boys, 19 girls) aged 6 to 12 years diagnosed with DLD, recorded with MD SONY MZ-N710 (fs = 44.1 kHz, 16-bit resolution), and 44 typically developing children (15 boys, 29 girls) aged 4 to 12 years, recorded with a SONY digital Dictaphone (fs = 16kHz, 16-bit resolution). Specifically, the dataset includes recordings of each letter of the alphabet. For the purposes of this study, MFCCs and their first and second derivatives were extracted from the recordings of the letter "A" to perform the classification task. Moreover, the mean and median frequencies of the power spectrum were computed, aiming to investigate eventual spectral distinctions between the two groups. The processing pipeline is shown in Figure 1.

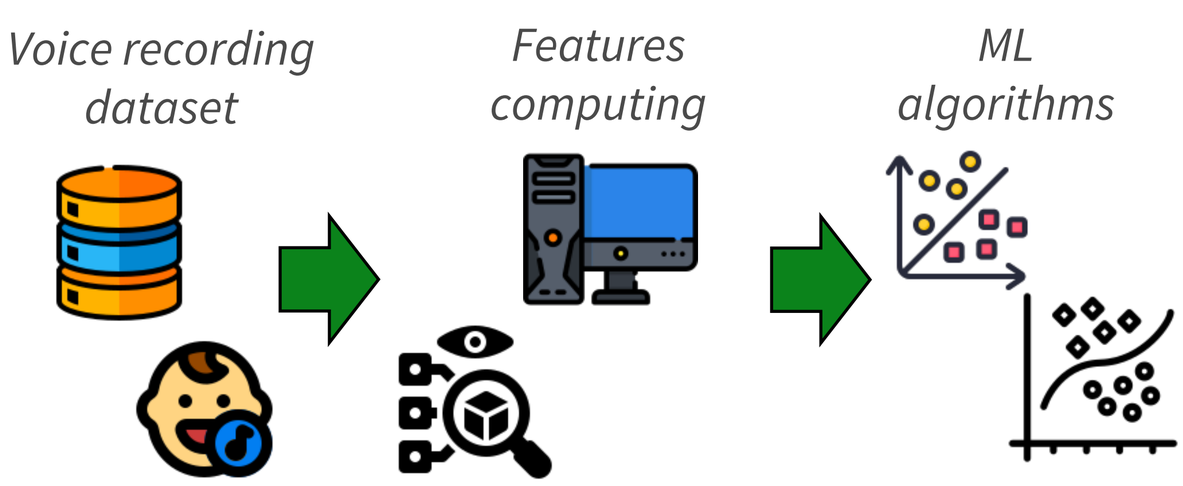

Fig. 1 Processing pipeline of the study

RESULTS & DISCUSSION

The best performance was achieved by employing a set of 15 features selected through the MRMR procedure, which resulted in a test accuracy of 96.7% and an AUC of 0.98, utilizing a Cubic SVM, as reported in Figures 2 and 3.

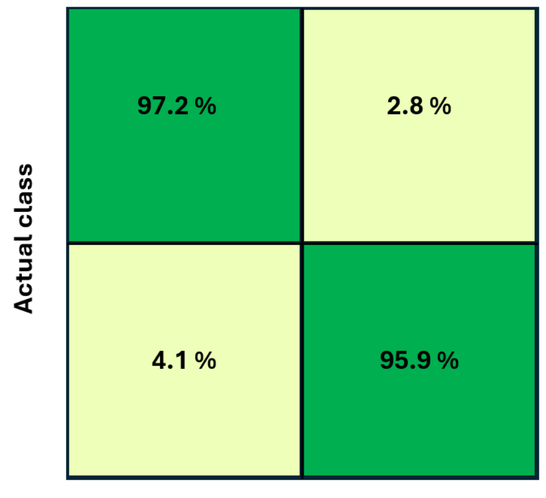

Fig. 2 Confusion matrix associated with the test set

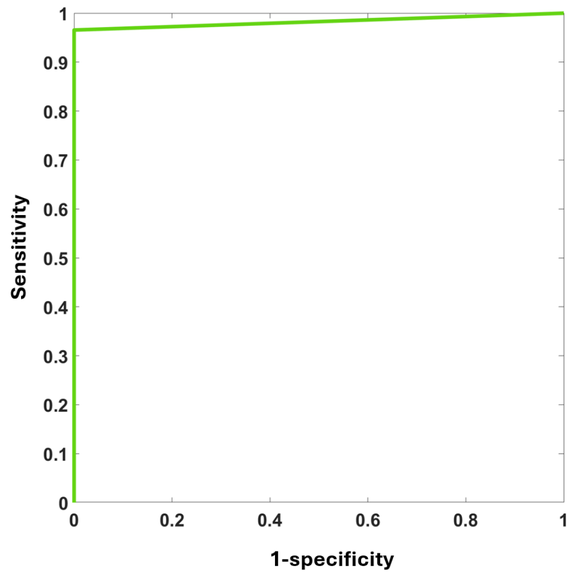

Fig. 3 ROC curve delivered by the model.

Moreover, a t-test assessed differences between the two groups concerning the mean frequency (p = 0.024) and the median frequency (p = 0.022).

CONCLUSION

This study demonstrated the feasibility of employing ML algorithms in diagnosing DLD through the analysis of MFCCs extracted from voice recordings. This approach could significantly enhance long-term outcomes for individuals affected by this disorder.

REFERENCES

1. Abdul, Z. K., & Al-Talabani, A. K. (2022). Mel frequency cepstral coefficient and its applications: A review. IEEE Access, 10, 122136-122158.
2. Grill, Pavel (2016). Speech databases of typical children and children with SLI. figshare. Dataset. https://doi.org/10.6084/m9.figshare.2360626.v2