

Stuttering Detection Application

Kowshik Siva Sai Motepalli*, Vamshiraghushimha Narasinga*, Harsha Pathuri, Hina Khan,
Sangeetha Mahesh, Ajish K. Abraham, Anil Kumar Vuppala

kowshik.siva@students.iiit.ac.in, narasinga.vamshi@research.iiit.ac.in,
harsha.pathuri@students.iiit.ac.in, hina8655@gmail.com, sangeethamahesh@aiishmysore.in,
ajish68@aiishmysore.in, anil.vuppala@iiit.ac.in

Abstract

Stuttering is a prevalent speech disorder that affects millions of people worldwide. In this Show and Tell presentation, we demonstrate a novel platform that takes speech samples in English and Kannada to detect and analyze stuttering in patients. The user-friendly interface includes demographic details and speech samples, generating comprehensive reports for different stuttering disfluencies. The platform has four different user types, providing full read-only access for admins and full write access for super admins.

Our platform provides valuable assistance for speech-language pathologists to evaluate speech samples. The proposed platform supports both live and recorded speech samples and presents a flexible approach to stuttering detection and analysis. Our research demonstrates the potential of technology to improve speech-language pathology for stuttering. Used F-score as a metric for evaluating the models for the stutter detection task.

Index Terms: stuttering, speech disorder, platform, recorded speech samples, syllable repetition, prolongation, blockage, comprehensive report

1. Introduction

Stuttering is a common speech disorder that affects approximately 1% of the world's population [1]. It is characterized by abnormal and continuous pauses in the normal flow of speech, and can be broadly categorized into different disfluencies, as shown in Table 1. Identifying and classifying stuttering is a challenging research problem that spans different domains, such as pathology, signal processing, and acoustics.

Detecting stuttering is important for providing timely and appropriate intervention to patients who stutter, and for preventing or minimizing the adverse effects of the disorder on their quality of life. However, the traditional method for evaluating stuttering severity by Speech Language Pathologists is manual speech recording analysis, which has several limitations. It is costly, time-consuming, and tedious and requires human expertise, which may introduce errors.

To overcome the limitations faced by speech-language pathologists (SLPs), we propose a novel platform that leverages the benefits of deep learning technologies. In the literature [2, 3], machine learning algorithms have shown promising results in identifying and quantifying stuttering patterns. Our platform uses syllable-level detection from recorded speech samples to analyze stuttering in patients, thereby reducing the burden on SLPs. It provides a user-friendly interface for recording demographic details and speech samples, supporting both real-time and pre-recorded file analysis for different types of stuttering disfluencies.

Welcome to your dashboard, me@superadmin!

You are logged in as SUPERADMIN.

DOCTORS PATIENTS



Figure 1: Super Admin Page

Abbreviation	Description
Blocks	Interruptions in the normal flow of speech
Prolongations	Lengthening of sounds or syllables
Part-word repetitions	Repeating part of a word
Word repetitions	Repeating a whole word
Phrase repetitions	Repeating a group of words
Mono-syllabic word repetitions	Repeating a word with one syllable

Table 1: Types of disfluencies in stuttering.

2. Application Overview

The Stuttering Detection Platform has a user-friendly interface that enables doctors, patients, and admins to navigate and interact with the system effectively. The platform has four types of users: patients, doctors, admins, and super admins. Patients can log in to view their test reports. Doctors can perform new tests, view test reports, and manage patient information. Admins have full read-only access to the application, Super-admins have full write access, as shown in Figure 1. This platform supports both English and Kannada languages.

The core of the Stuttering Detection Platform is the new test creation feature, which allows doctors to enter patient demographics such as Case Name, Case Number, Age, Email, Family History, Occupation, Address, Nature of the Problem, Duration and an option to select model specific to English or Kannada language.. The speech samples contain spontaneous sentences and paragraphs, as shown in Figure 3. The doctor can add new questions from the existing database as needed. The platform also enables doctors to record patient responses or upload pre-recorded files. The database contains many spontaneous sentences with questions like "What is your name?", and "What are your hobbies?", and paragraphs taken from a short story "The Wonderful Story of Henry Sugar and Six More". The system uses the collected speech samples to generate a report that includes predicted syllables and the number of syllables involved

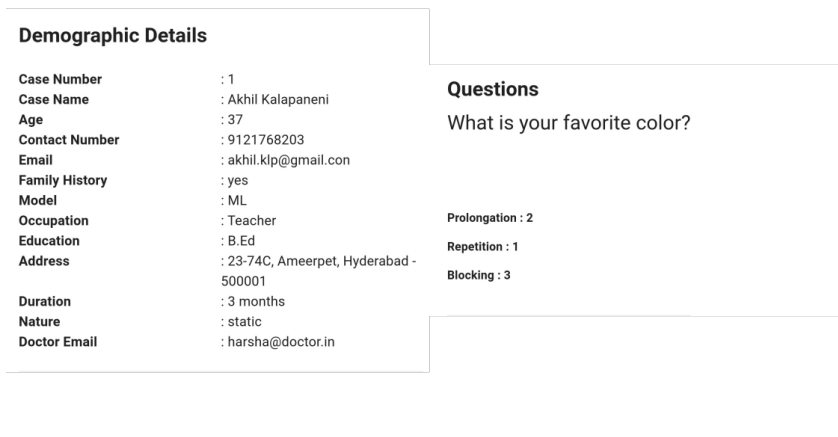


Figure 2: Report

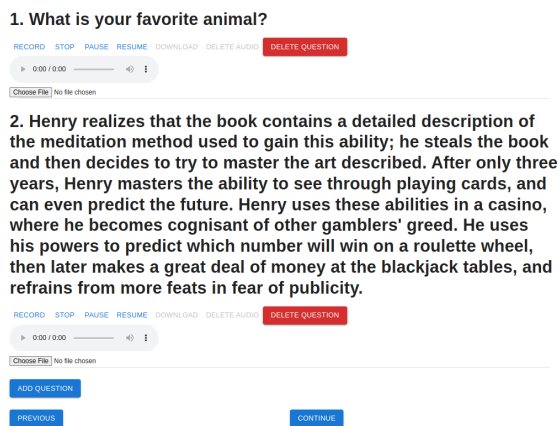


Figure 3: Questions

in blocking, prolongation, and repetition, as shown in Figure 2. We created several machine learning and deep learning models such as SVM, LSTM, BiLSTM [4] for the task of stutter detection and models are trained on SEP-28k database [5] for English language and for Kannada, we have collected few samples. F-score is used as a metric for evaluating the model performance. Models are trained on standard handcrafted features such as MFCC(Mel Frequency Cepstral Coefficients) and Zero Time windowing Cepstral Coefficients(ZTWCC) and results were tabulated in Table 2, where we can observe that ZTWCC features with BiLSTM is performing better compared to rest as it captures high resolution features better than MFCC's and it is used in the final version of the platform.

3. Conclusions

The Stuttering Detection Platform is a web-based system that analyzes stuttering instances in patients in real-time by recording the data of the patients in real time and using the platform to generate the results. This enables speech pathologists to detect stuttering more accurately and efficiently. The report generated from this platform can provide critical insights into the nature

Table 2: F-Score Performance of Stuttering Classification by Different Classifiers

Classifier	Feature type	Stutter type		
		B	W	P
SVM	MFCC	0.6210	0.6126	0.6514
	ZTWCC	0.6008	0.5783	0.6261
LSTM	MFCC	0.5585	0.6929	0.7307
	ZTWCC	0.68966	0.7061	0.7375
BI-LSTM	MFCC	0.7129	0.6977	0.7242
	ZTWCC	0.7223	0.7045	0.7359

Abbreviations: B = Blocking, P = Prolongation, W = Word Repetitions

of the disorder. The Stuttering Detection Platform holds great promise for improving the quality of care for patients who stutter. It represents a significant advancement in stuttering detection and can become an essential tool for healthcare professionals.

4. References

- [1] A. Smith and C. Weber, "How stuttering develops: The multifactorial dynamic pathways theory," *Journal of Speech, Language, and Hearing Research*, vol. 60, no. 9, pp. 2483–2505, 2017.
- [2] T. Kourkounakis, A. Hajavi, and A. Etemad, "Detecting multiple speech disfluencies using a deep residual network with bidirectional long short-term memory," in *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2020, pp. 6089–6093.
- [3] S. A. Sheikh, M. Sahidullah, F. Hirsch, and S. Ouni, "Stutter-net: Stuttering detection using time delay neural network," in *2021 29th European Signal Processing Conference (EUSIPCO)*. IEEE, 2021, pp. 426–430.
- [4] I. Esmaili, N. J. Dabanloo, and M. Vali, "An automatic prolongation detection approach in continuous speech with robustness against speaking rate variations," *Journal of medical signals and sensors*, vol. 7, no. 1, p. 1, 2017.
- [5] C. Lea, V. Mitra, A. Joshi, S. Kajarekar, and J. P. Bigham, "Sep-28k: A dataset for stuttering event detection from podcasts with people who stutter," in *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2021, pp. 6798–6802.