

PRODUCT REPORT

2024 AI CHALLENGE- STUTTERING THERAPY SYSTEM

PRODUCT

- **Problem**

Stuttering is a speech disorder characterized by interruptions in the flow of speech. This speech disorder is mainly divided into three types; repetitions, prolongations, and blocks. Stuttering affects individuals of all ages and can have a significant impact on their communication abilities and self-confidence.

- **Current Solutions**

Current speech therapy methods primarily target enhancing speech fluency through techniques like breathing exercises and deliberate speech. However, these approaches can be inaccessible and costly. Hence, there's a demand for an innovative, personalized solution to better aid individuals dealing with stuttering.

- **Our Solution**

Our system offers a comprehensive approach to address this issue. Here's a brief overview of the functionalities:

1. **Stuttering Detection:** Utilizing cutting-edge speech analysis algorithms, our system accurately detects instances of stuttering in individuals' speech patterns.
2. **Classification into Main Types:** Identified instances of stuttering are classified into main types such as prolongation, repetition, and blocks.
3. **Personalized Therapy Provision:** For each identified type of stuttering, therapy resources in the form of audio recordings are provided.

- **Target Audience**

Our system is for individuals of all ages facing stuttering or speech fluency issues, with a specific emphasis on those who stutter, regardless of age. Additionally, speech-language therapists and professionals assisting individuals with stuttering can also find value in our system.

ENGINEERING

High-Level Design:

Our stuttering detection system operates on a two-fold approach: Audio classification and Therapy generation. Using the DistilHubert transformer model, audio inputs are classified into distinct stuttering types. This process involved preprocessing the data, training the model on our own dataset, and optimizing its performance through fine-tuning. The system's architecture ensures efficient classification of stuttering types, providing users with valuable insights into their speech patterns.

Role of Generative AI (MMS Model Text-to-Speech Therapy Recommendation):

We used the Massively Multilingual Speech (MMS) English Text-to-Speech model for therapy generation based on the identified stuttering type. By using it as an API in our system, users can receive therapy in the form of audios.

Technologies Used:

- DistilHubert (transformer model)
- Massively Multilingual Speech (MMS) (text-to-speech model)
- Hugging Face Transformers library
- Gradio library
- TensorFlow
- PyTorch

RESULTS AND ANALYSIS

In this section, we present the results of our stuttering detection model which was fine tuned on the **Distil Hubert** model using our own dataset.

- Evaluation Approach: For evaluating the model's performance, we computed various metrics, including accuracy, on an independent test set.
- Loss: The model achieved a loss value of 0.5717, indicating the variation between the predicted values and the ground truth labels during training.
- Accuracy: With an accuracy score of 0.9024, our model correctly classified approximately 90.24% of the samples in the evaluation set.

FUTURE PROSPECTS

1. Multimodal Inputs: Integrate additional features such as facial expressions or physiological signals to improve the accuracy and robustness of stutter detection.
2. Personalized Therapy: Tailor therapy recommendations to individual preferences and needs.
3. Real-Time Feedback: Provide immediate support during speech interactions for improved fluency.
4. Accessibility and Scalability: Optimize resources and support multiple languages for broader user reach.
5. Collaboration with Healthcare Professionals: Partner with experts to validate effectiveness and ensure alignment with established practices.