CONTENTS

Motivation

Over 5% of the world's population(430 million people) require rehabilitation to address their 'disabling' hearing loss (432 million adults and 34 million children).
They more often than not have imperfect development of speech & language
It is estimated that by 2050 over 700 million people – or one in every ten people – will have disabling hearing loss.

People with hearing disabilities regularly face difficulties while having conversations on different occasions.
A hearing aid is a wonderful device, helpful for those people whose hearing disabilities can be cured using it and those who can afford it. But a lot of people's disabilities cannot be cured using a hearing aid or they cannot afford it. For such people, the only option is either communicating using sign language and using lip-reading or a combination of both.

Sign language has evolved manifolds over the years and it allows for effective communication but the huge catch is not many people in the world sign language primarily because it is not that easy to learn and can take years to gain even a passable proficiency.

Now the benefit of lip reading is that the other person does not need to learn sign language to communicate and they can do so by correctly enunciating their speech.

Our App focuses on helping such people by features like "Lip Reading Tutorial", "Sign Language Tutorial". These features help them practice and improvise their abilities to understand whatever the other person is saying while having a real conversation.

On the other hand, "Speech Assistance" features help us know how fast or slow we are talking so that we can change the speed accordingly to give a better presentation.

While watching videos a person with a hearing disability often can't understand what the person is trying to convey because lip-reading at that pace is almost next to impossible. The feature "Video Transcript" generates auto-subtitles and can help them better understand the context of the video.

User Workflow Diagram
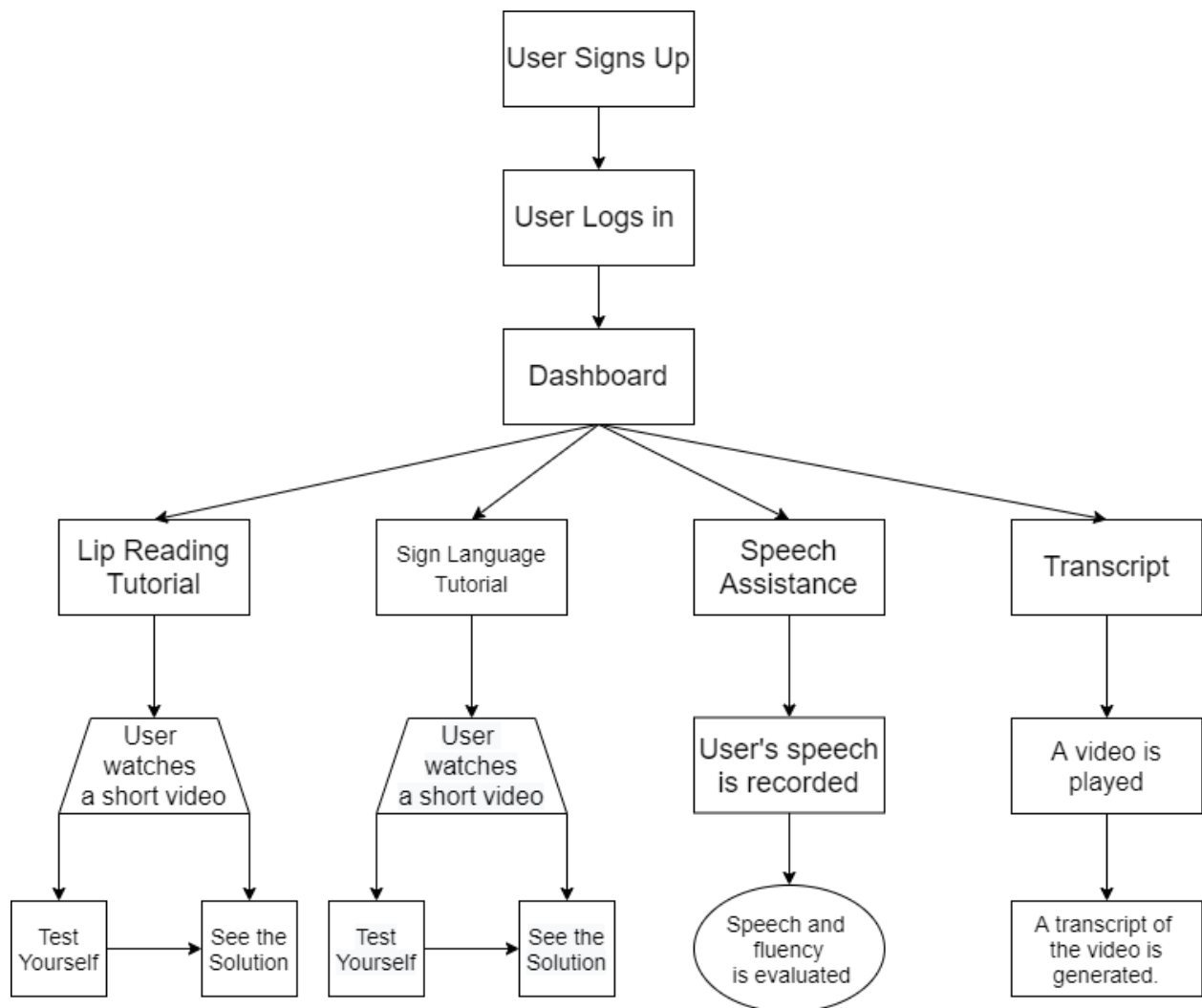
As can be seen from the above diagram first a user needs to register and then after a successful registration they log in. Once the user logs in they are greeted by a dashboard and from here they can navigate to the feature that they would like to try from among Lip Reading Tutorial, Sign Language Tutorial, Speech Assistance, and Transcript.

Features

Lip Reading Tutorial

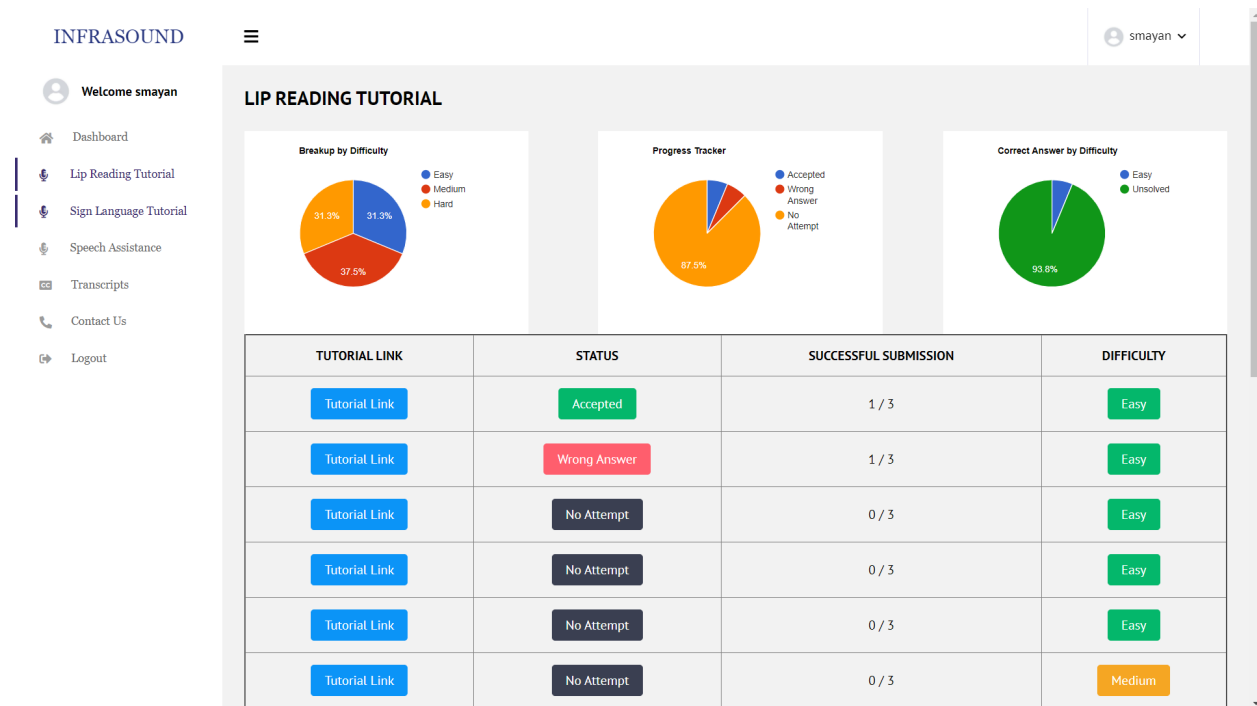This feature allows the user to evaluate their lip-reading capabilities.

The image below shows three graphs that keep a track of the user's progress.

In the table, the first column contains the tutorial(exercise) link.

The second column contains the status of whether the tutorial is attempted, correctly answered, or wrongly answered.
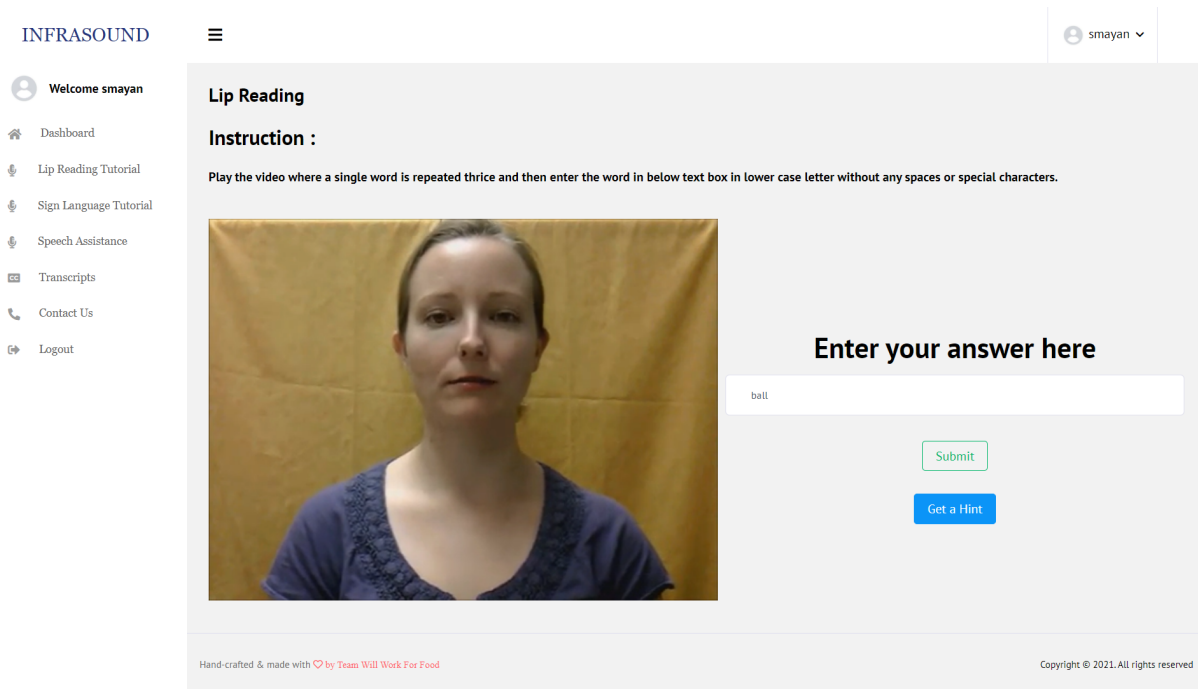
The third column gives a comparison between other users depicting how many of the total users have correctly answered the question.

And the last column depicts the difficulty between easy, medium, and difficult.


The graphs shown here are dynamic meaning, they change as and when the user correctly or incorrectly answers a question.

The user is prompted to first play the video where the tutor clearly enunciates a single word thrice and then the user is prompted to enter their answer in the text box.

In case the user is unable to answer correctly after many attempts, they can look for the hint using the **Get a Hint** button
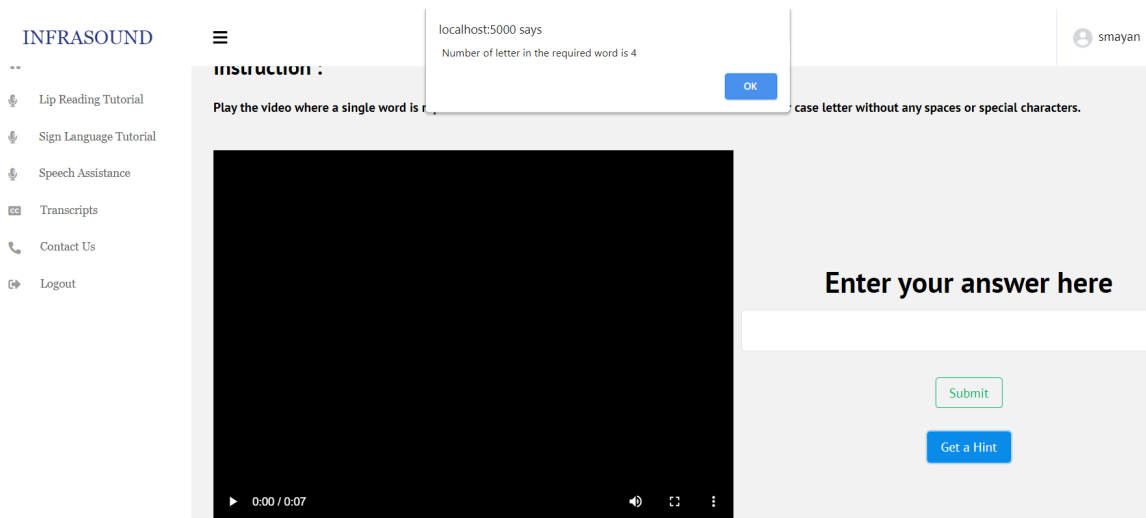
Sign Language Tutorial

This feature allows the user to evaluate their sign language capabilities and is pretty much the same as the previous lip-reading tutorials.

The image below shows three graphs that keep a track of the user's progress.

In the table, the first column contains the tutorial(exercise) link.

The second column contains the status of whether the tutorial is attempted, correctly answered, or wrongly answered.

The third column gives a comparison between other users depicting how many of the total users have correctly answered the question.

And the last column depicts the difficulty between easy, medium, and difficult.
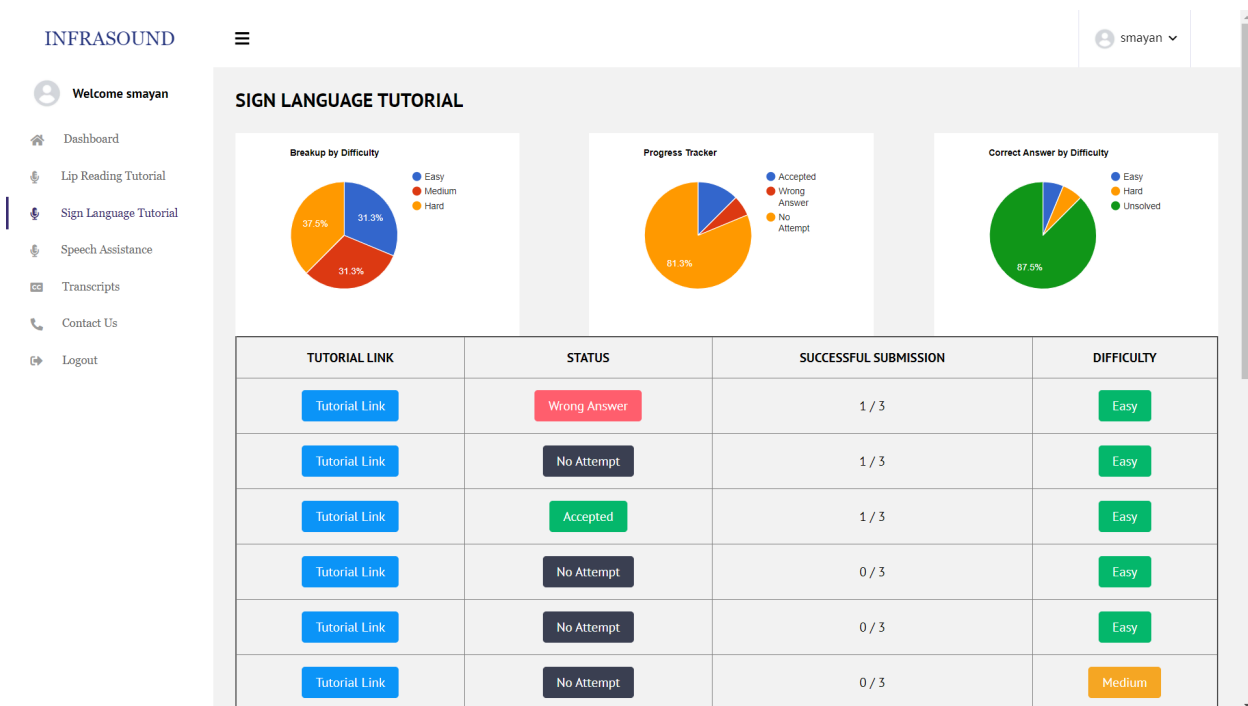
The graphs shown here have dynamic meaning, they change as and when the user correctly or incorrectly answers a question.

The user is prompted to first play the video where the tutor signs clearly a single word and then the user is prompted to enter their answer in the text box.
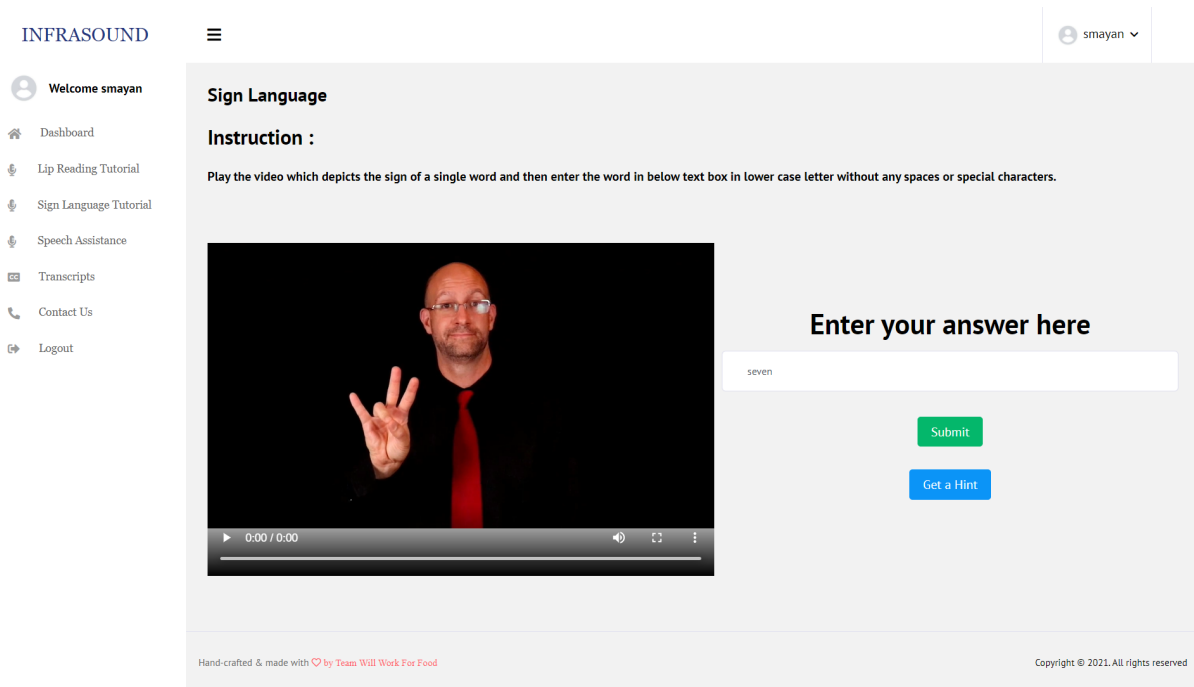
In case the user is unable to answer correctly after many attempts, they can look for the hint using the **Get a Hint** button
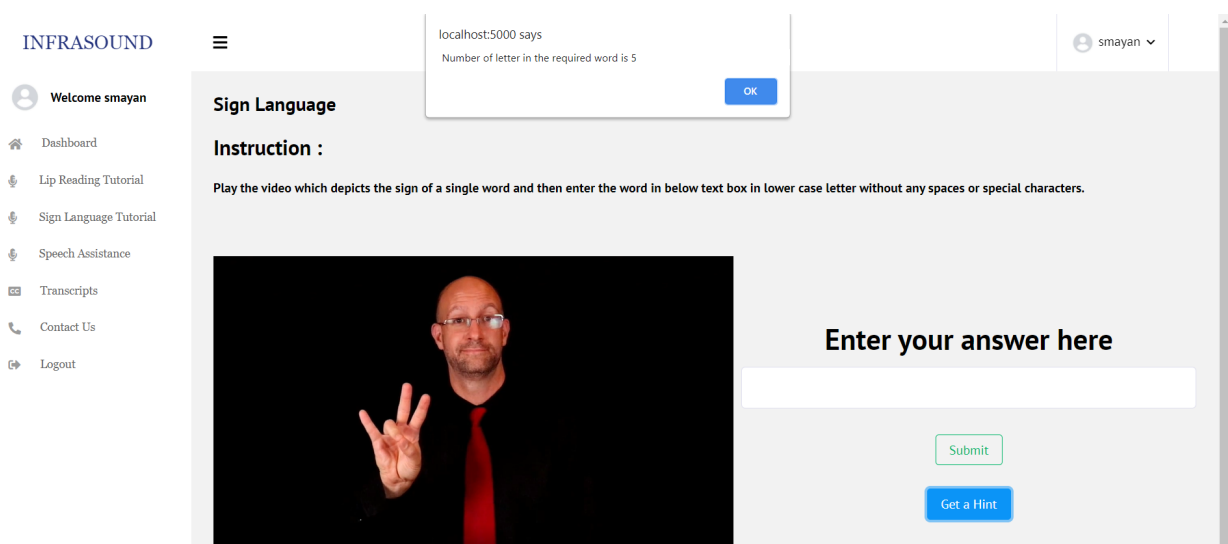
Speech Assistance

This feature allows the user to evaluate the characteristics of their speech like loudness and speed and helps them improve that over time using this feature.

The image below shows the portal. In the backend, there's an integrated speech analyzing engine and a speech-to-text engine that work in tandem to generate the text from the speech and the speed with which you are speaking.

The speech-to-text helps users gauge the correctness of their pronunciation and the speed helps them determine whether to try to speed up while speaking with other individuals.

The user clicks on the Take a test now button and then they are prompted to speak as shown below.
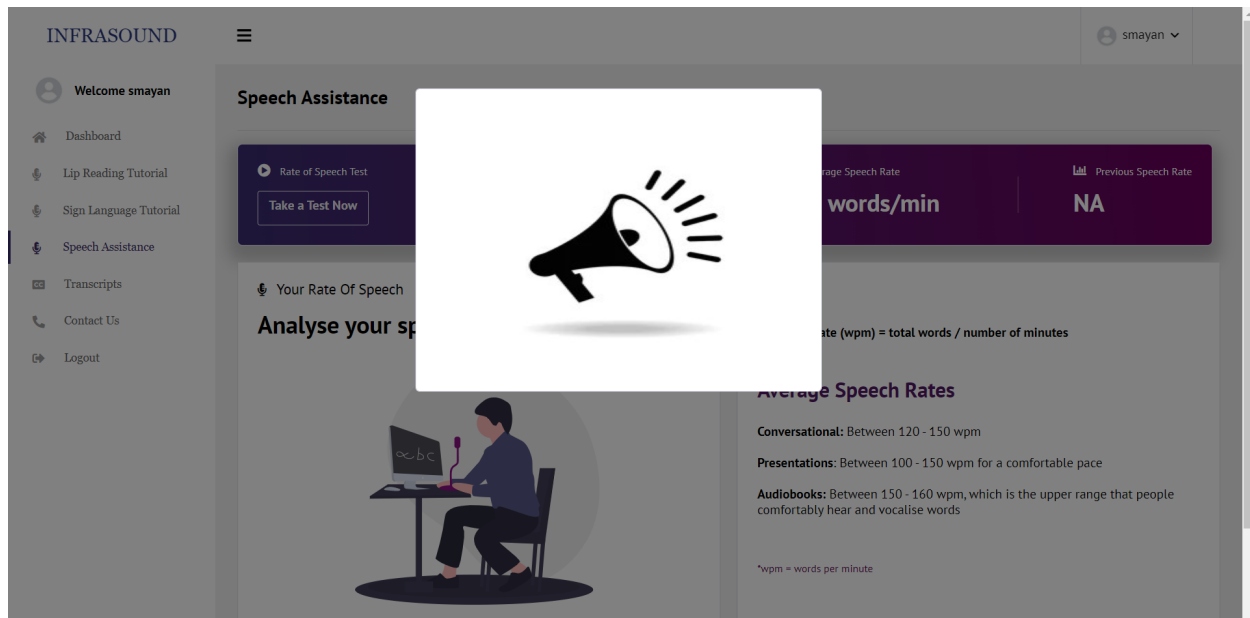
Once the dialog box showing the loudspeaker closes, the output is generated at the bottom as shown below.
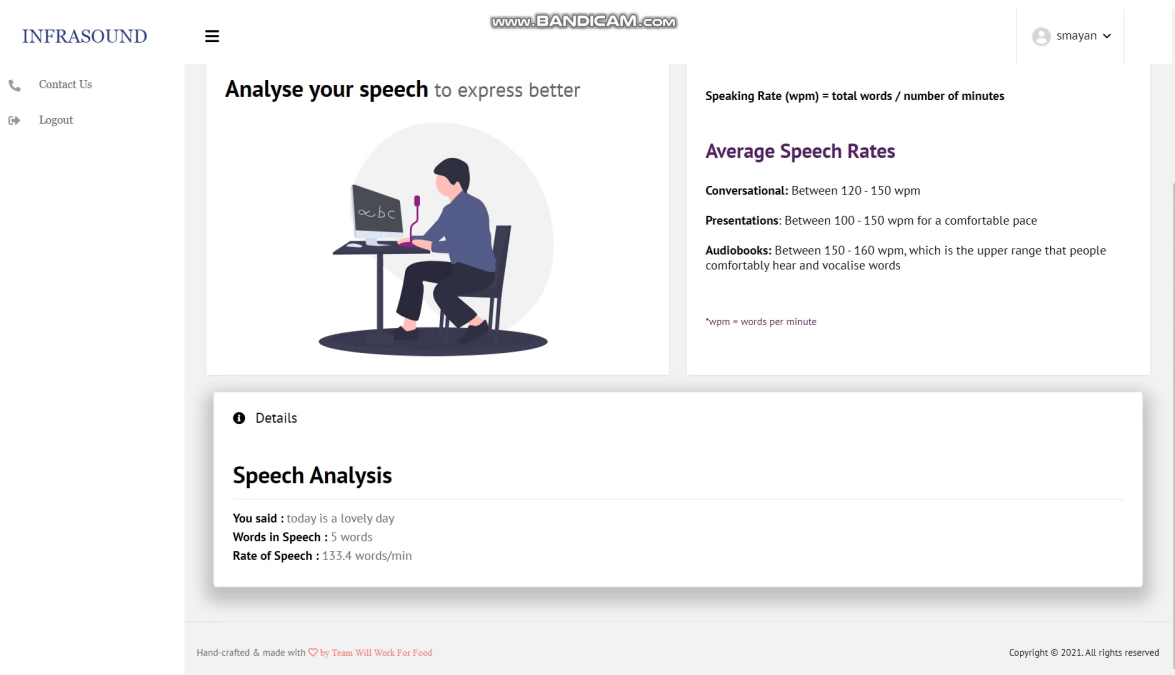
<u>Transcript</u>

Here a video is played and there's an audio-to-text converter integrated that tries to generate a transcript of the video.

Although we tried to make it as real-time as possible, there's a lag of about 6-7 seconds after the audio is heard.

Future Work and Enhancements

Lip Reading and Sign Language

In the lip reading part the tutorial only focuses on single words, but it can be extended to a sentence to provide advanced exercises and the same can be implemented for sign language.

Speech Assistance

There can be added functionality where a text to speech engine first converts the text entered to speech and then a human face that enunciates clearly how to speak that sentence. This will help users imitate the same sounds and thus improve their pronunciation capabilities.

Transcript

The major problem with the transcript feature was the inefficiency of the speech-to-text engine as it's a multi-step process of first extracting the audio from the video and then parsing the audio to generate text. This currently makes the text lag 6-7 seconds behind the video but it can be improved significantly!

How to Run the application

Prerequisites:

Make sure you have Python 3.7 or higher installed on your system to run the application.

Also install venv library to install the dependencies in a virtual environment using the command:

pip install venv

Steps:

1. First download the zipped folder **Code.zip** and then unzip it.
2. Open a Command Prompt/Terminal and navigate to the folder where you have unzipped it and enter that directory.
3. Create a virtual environment using the following commands:
   virtualenv t4sne

   Then to activate the environment type:
   For windows:
   t4sne\Scrips\activate

   For Mac/Linux:
   source t4sne/bin/activate
4. Now type the following to install all the dependencies:

   pip install -r requirements.txt

   Wait for it to install everything.
5. Next type:
   pipwin install pyaudio
6. Now finally run:
   python application.py

   And then open any browser and type in localhost:5000 or 127.0.0.1:500