**Objective:**

* **Magasool** launched Uzhavar Shakti in partnership with https://gramvaani.org/ in March - April 2021 with 1300 farmers. Uzhavar Shakti (Farmers' Power) is a localized, remote extension and market access system for farmers and farm workers.Content is curated from existing services, Government and private information portals and village resident experiences. 9000+ farmers now subscribe to this service.
* In the Uzhaivar Sakthi radio program, farmers express their concerns or issues as requests. These requests are recorded in the Uzhaivar Sakthi system. Subsequently, a representative from Uzhaivar Sakthi reviews the requests, categorizes them into specific sections, and forwards them to the relevant officials responsible for addressing those issues
* To streamline the process, this project aims to develop automatic speech-to-text recognition that categorizes and forwards the requests automatically.

**Our Short term Goals:**

1. Collect the recordings of calls

2. Convert them to text

3. Match the important keywords

4. Make two categories of urgent and not urgent calls.

**Choosing API’s:** For this we will start by using an open source tool called **BHASHINI**. This gives API’s for audio to text, translation, and other things.

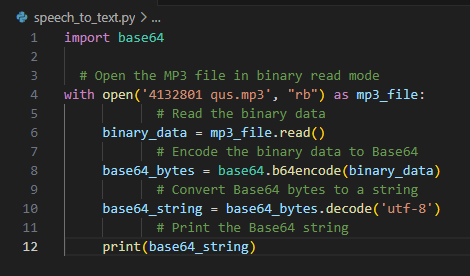
**Programming Language: Python** platform easily understand for beginners.

1. Create a python project with one file with the name speech\_to\_text.py
2. Collect sample audio files of the calls (at least 20) and put them in the same folder as your python file
3. BHASHINI only accepts the audio file in base64 format as the input. So first we must convert the audio files to base 64. You have to use the python base64 package



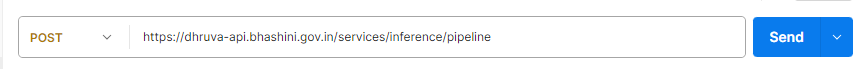
1. In the speech\_to\_text.py file write a function called convert\_mp3\_to\_base64. This function will take an mp3 file as input and convert it to a base64 string. We will start by doing this for a single file

* First open any **one of the** MP3 files in binary read mode
* While the file is open, read the file and store it in a variable
* Encode the binary data to Base64
* Convert the Base64 bytes to a string with 'utf-8' encoding
* Print the string

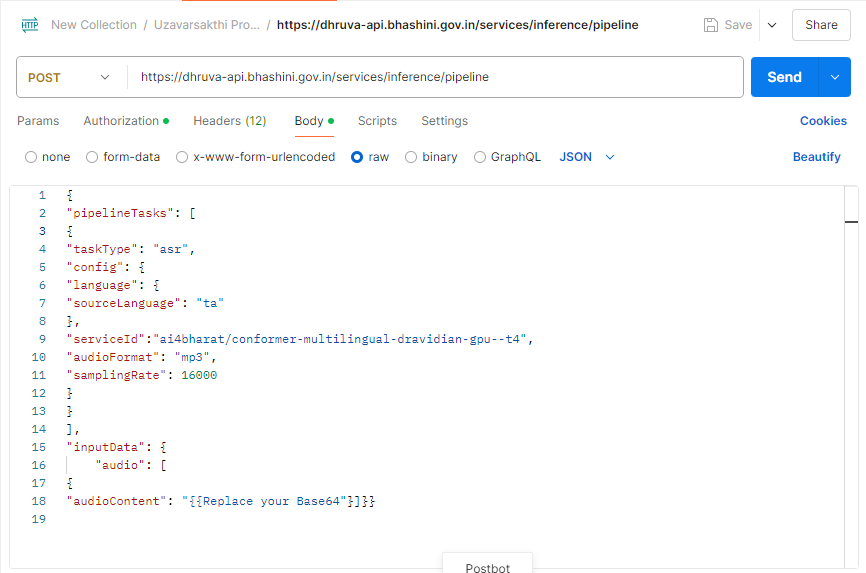


**API Documentation:**

1. Now we will call the BHASHINI API.
2. First we will try it from **postman** it integrate python code and BHASHINI
3. Download **postman** on your computer
4. Call an API by:
   * Setting the request type to POST
   * Setting the URL to <https://dhruva-api.bhashini.gov.in/services/inference/pipeline>



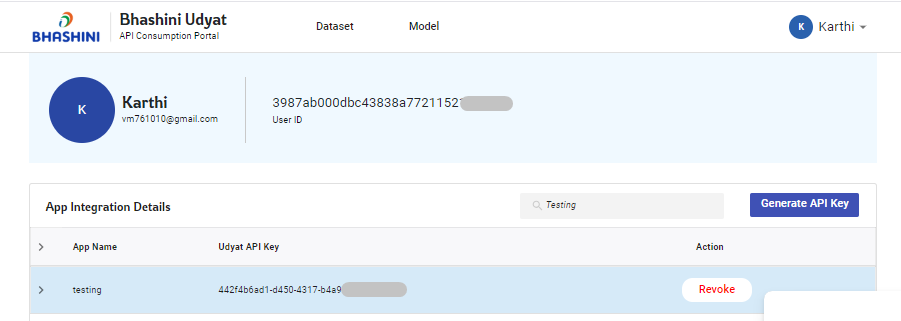
* + Set the request body to this given below. Replace {{THIS\_IS\_YOUR\_BASE\_64\_STRING}} with the base64 string your printed in the last step



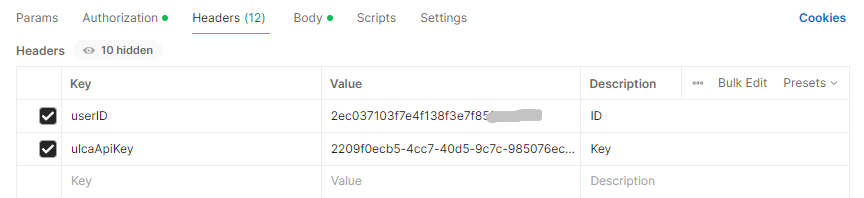
1. In the post request successfully make we need API key, User ID.
2. **How to get Keys**:

* Overall understanding of the API Calls, In this link helps understand the concept of API.
  + Reference Link: <https://bhashini.gitbook.io/bhashini-apis>
* **Account creation:** This page will help the integrator to get themselves onboarded on BHASHINI and get the required API Keys and User ID.
  + Registration: [https://bhashini.gov.in/ulca/user/register#](https://bhashini.gov.in/ulca/user/register)
* **API Key Generate:** Integrators will be able to a create the API Key using **Generate**

Button under their. So we got API Keys and User ID.

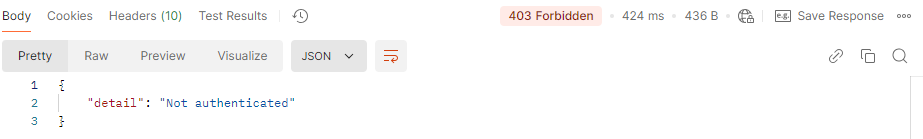


1. Using Obtaining Keys in postman, **ULCA API’s Keys** and **User ID** using postman of header section, Add extra needed names.



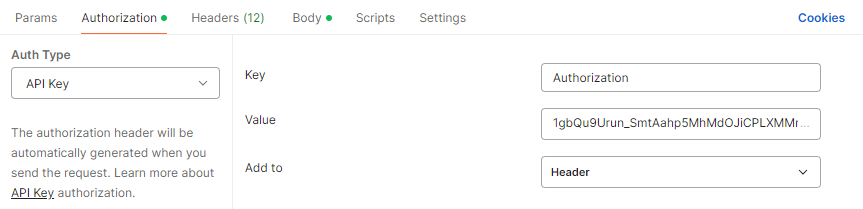
* I think all the process are completed, So I start make the post request to server then click send button.

**Response i got** **403 forbidden** detail **(“Not Authentication”) Error shows**



|  |
| --- |
| **Problem 1;**     * First time facing big problem because I don’t know what **Error a**nd **next what can I do.** I am struggling this part of Error. * So i started search solution of the problem. API related websites and YouTube channels   Know one can tell the exact solution of my problem.   * One day I got opportunity chance to meet one person. He working from IT API’s development field. I say my problem and struggling of the project.   **Solution:**     * He tell about one answer we need authorization key of BHASHINI from Official * How we Get: Compose the mail request to BHASHINI about purpose of project so need authorization. * **Finally Obtain Authorization Key** |

1. Authorization Key using authorization Section of postman



1. Last step send the request, see the results.



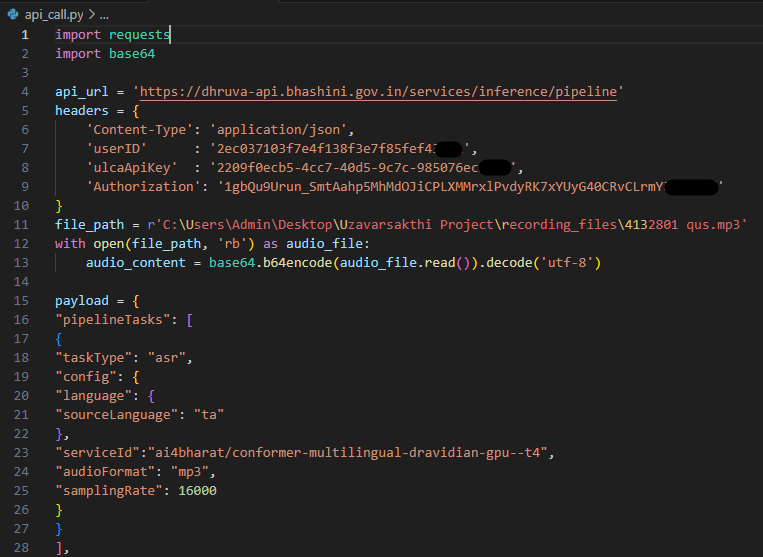
1. **finally we got expected results.**
2. Check the output to see if it is correct.

If it is not working or not correct, try with 3-4 different files. Note the results.

1. Now call the same API directly from the python code.
   1. First, google to understand which library to use for calling API from python code

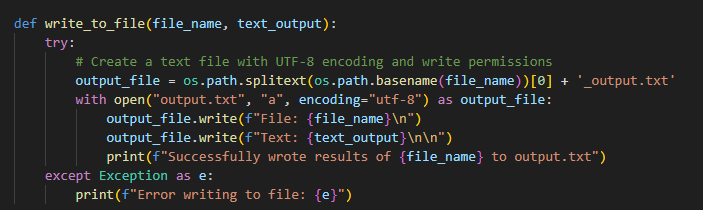


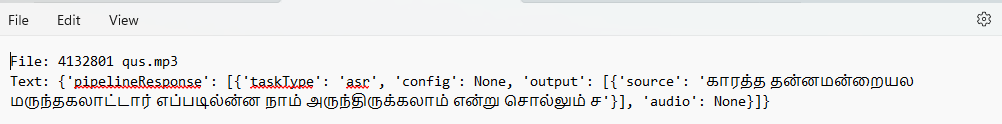
* 1. Set up the API correctly as a POST request with the correct URL, headers, and request body
  2. Print the response body
  3. Add try and catch statements so catch and debug any errors



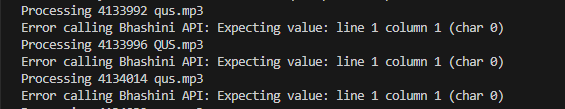
1. Create a file that has the name of the recording file and the text of the call
   1. Using python create and open a .txt file with write permissions
   2. Write the name of the recording file and the text outputted by the Bhashano API to that file
   3. Make sure the file is easy to read



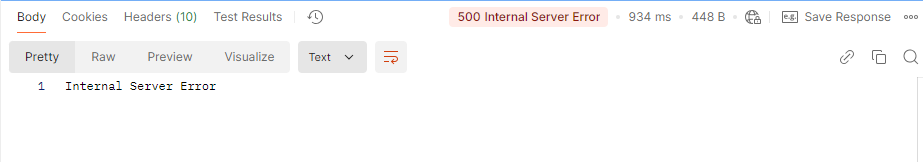


**Output:**

* **Most of mp3 audio files failed will be transcript output shows**



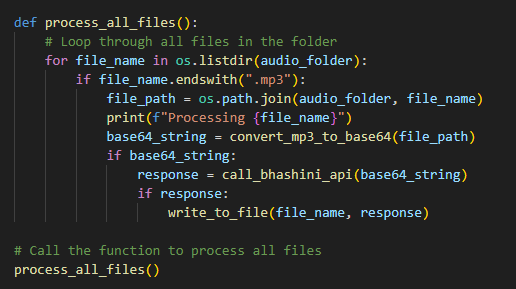
* **Postman API’s shows 500 Internal Server Error**



* **I collect multiple amount of mp3 audio files for testing purpose most of mp3 files transcript will be errors shows.**

|  |
| --- |
| **Problem 2:**   * **Above 25sec mp3 audio files cannot transcript from BHASHINI.** * **I research about ASR (automatic speech recognition) some of the ASR transcript 25sec audio files for capacity at a time**   **Solution:**   * **We can split the audio file as 20 sec as input and results are something looking good** |

1. Now we will do all the steps from last week in a loop for all the sample files.
2. In the python code, create a loop to open the audio file, convert to base64, call the BHASHINI API and write to the file.
3. The part to think about is how to add error handling so that the code does not stop at one step if there is an error. It should reach the end.



**Output:**

* File: 4132801 qus.mp3 **(BHASHINI OUTPUT)**
* Text: {'pipelineResponse': [{'taskType': 'asr', 'config': None, 'output': [{'source': 'காரத்த தன்னமன்றையல மருந்தகலாட்டார் எப்படில்ன்ன நாம் அருந்திருக்கலாம் என்று சொல்லும் ச'}], 'audio': None}]}

|  |
| --- |
| **ACTUAL OUTPUT:**  தென்னை மரத்துக்கு என்ன என்ன மருந்து வெக்கலம் சார், எப்டி வெக்கலம் சொல்லுங்க சார் |

* File: 4132874 qus.mp3 **(BHASHINI OUTPUT)**
* Text: {'pipelineResponse': [{'taskType': 'asr', 'config': None, 'output': [{'source': 'சந்நபடு அயாான கத்ரியாபடள சத்தகத்தி ஆயவுதியதி மருந்தடிய நம்ம்சடிவ'}], 'audio': None}]}

|  |
| --- |
| **OUTPUT:** அய்யா நா கத்திரிக்காய் பொடன் எல்லாம் சொத்தை  கத்திரிக்காய் அதுக்கு மருந்து அடிகணும் என்ன மருந்து அடிக்கலாம் |

**Conclusion:**

* The output from **Bhashini** contains many unclear words. This could be due to several reasons, such as background noise or low voice accuracy. I didn’t get the results I expected. So, this problem couldn't be solved using **Bhashini*.***
* I plan to try a few more APIs next for speech-to-text conversion. The results will become clear only after testing. The APIs I intend to explore are:

1. **Google Cloud Speech-to-Text**
2. **Microsoft Azure Speech Service**
3. **Amazon Transcribe**
4. **ChatGPT by OpenAI**