# Prerequisites:
1. Have python installed on your computer
2. Have a GitHub account

# Setup:
1. **Download the Git Repository:**
   https://github.com/AmoghBindal/MeetingBuddy

2. **Download the following libraries of python:**
   - replicate
   - sentence-transformers
   - opencv-python
   - pyautogui
   - pyaudio
   - nltk
   - SpeechRecognition
   - vosk

   **process:**
   In the command prompt run the following command:
   > **pip install {library_name}**

3. **Download the following transcription model:**
   https://alphacephei.com/vosk/models
   Extract the zip file into the folder where you have cloned the repository and
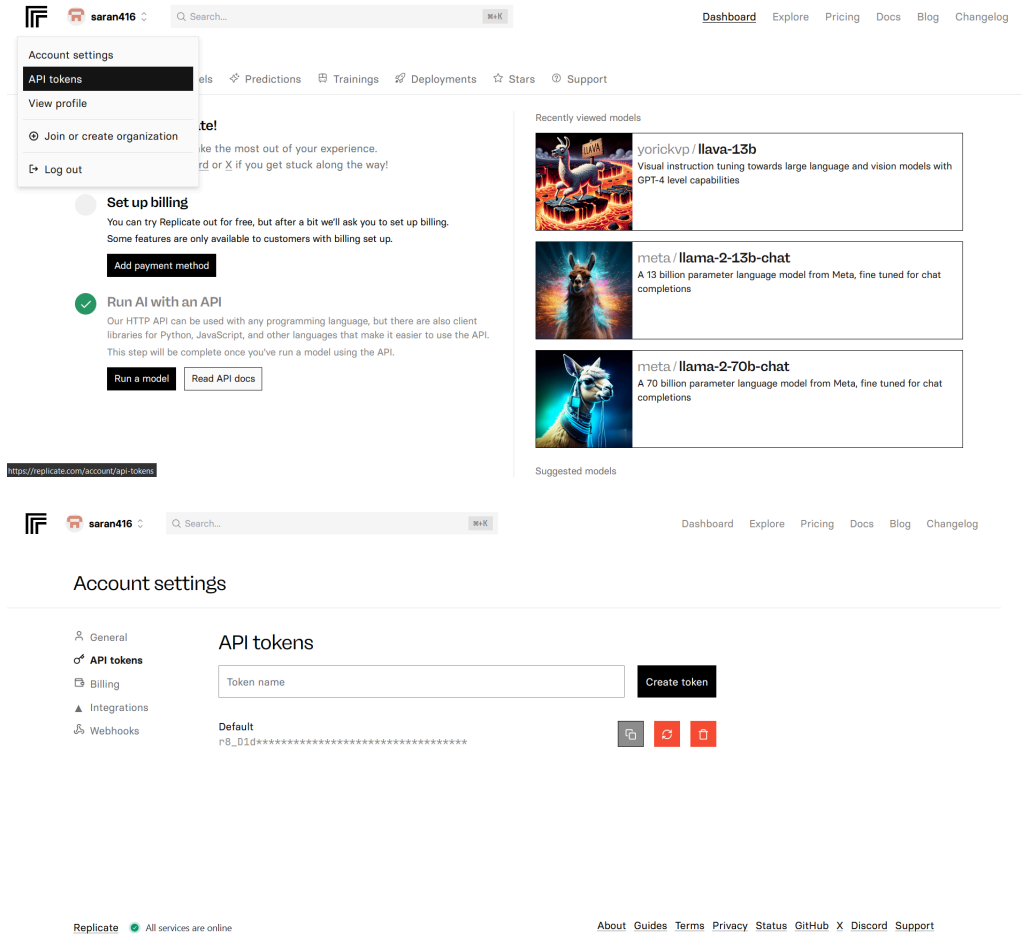   **rename it as "model".**

**Note:**
- The model we have used can also be replaced with whatever model you like based on the accent, language and precision required(more precision might require greater space and computing power).
- Model used by us in the demo is: **vosk-model-en-in-0.5.**

4. **Using llava API from replicate ai:**
   https://replicate.com/yorickvp/llava-13b

- Sign in with your github profile.
- On the top-left click on your profile name and then click on API tokens.
- Copy your Default API token.
- Now in the file **'llava.py'** replace "**{YOUR_REPLICATE_API_TOKEN}**" with the copied text.

```python
import replicate
import os
import base64
import ctypes  # An included library with Python install.

os.environ["REPLICATE_API_TOKEN"] = "{YOUR_REPLICATE_API_TOKEN}"

print(os.environ.get("REPLICATE_API_TOKEN"))

#This Function Takes text and image and prints response
def genresponse(text, imgurl):
    binary_fc       = open(imgurl, 'rb').read()  # fc aka file_content
    base64_utf8_str = base64.b64encode(binary_fc).decode('utf-8')

    ext     = imgurl.split('.')[-1]
    dataurl = f'data:image/{ext};base64,{base64_utf8_str}'

    output = replicate.run(
        "yorickvp/llava-13b:b5f6212d032508382d61ff00469ddda3e32fd8a0e75dc39d8a4191bb742157fb",
        input={
            "image": dataurl,
```

## 5. Change the name:

In the "Main.py" file replace {YOUR_NAME} with your name.

```python
import speech_recognition as sr
import NLP
import pyautogui
import cv2
import numpy as np
from llava import genresponse
import winsound

frequency = 2500  # Set Frequency To 2500 Hertz
duration = 2000   # Set Duration To 1000 ms == 1 second
personname = "{YOUR_NAME}"

#initialize recognizer
r = sr.Recognizer()

def record_text():
    #loop in case of erros
    while(1):
        try:
            #use the microphone as source for input
            with sr.Microphone() as source1:
                r.adjust_for_ambient_noise(source1, duration=0.2)
                print("Listening...")
                audio1 = r.listen(source1,0,8)
                print("Converting to Text....")
                MyText = r.recognize_vosk(audio1)

                return MyText

        except sr.RequestError as e:
            print(f"Could no request results {e}")

        except sr.UnknownValueError:
            print("unknown value occured")
```
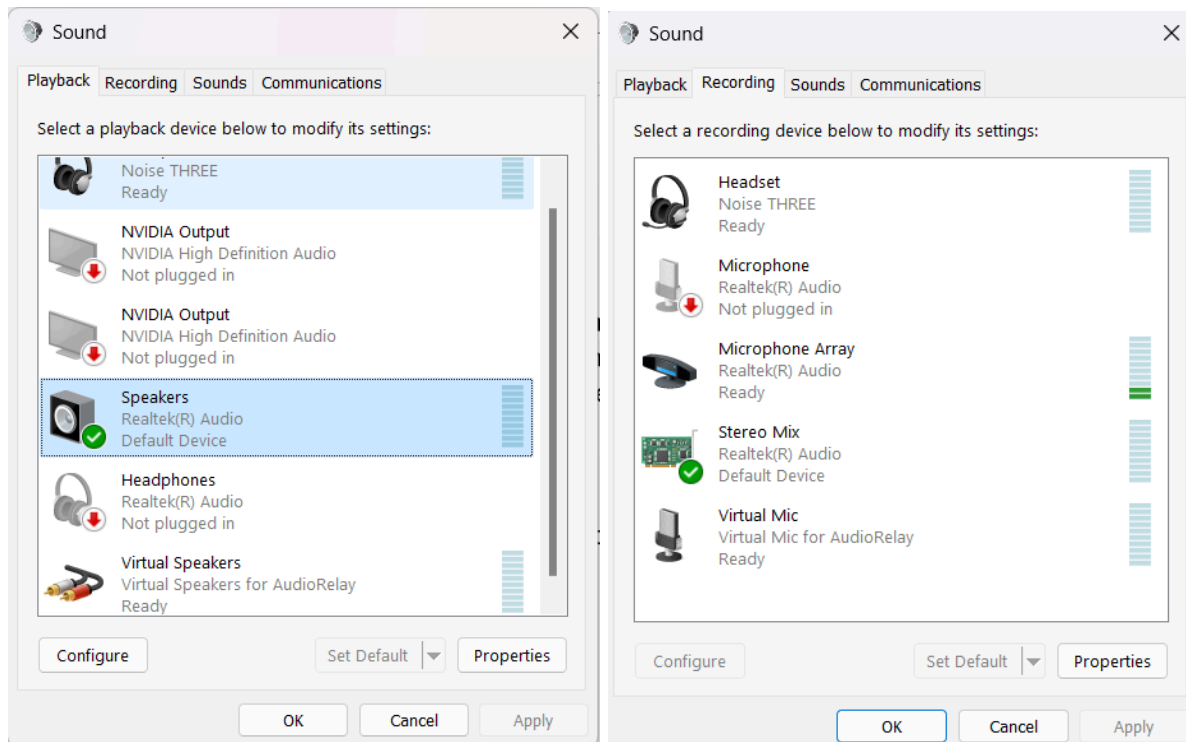
## 6.Changing default audio-input:

- Open windows run program [⊞ **Win + R**] and run **mmsys.cpl**
- Go to the **Playback** section in the top left corner and make sure the default device is **Device Speakers**(while running the meeting make sure audio is playing on the device speakers).



- Go to the **Recordings** section.
- Change Default device to **Stereo Mix.**
- When joining the video meeting make sure your Audio Input device is a Microphone and not Default settings since we changed the Default to Stereo Mix.
- When you are done using the application make sure to change the sound settings to previously set Default Devices.