



JULY 2024



# USER MANUAL - VIGOR



## VIGOR

### VIGOR LOGIN

- 1 VIGOR LOGIN Page
- 2 VIGOR Dashboard
- 3 VIGOR Application – (TRAINING)
- 4 VIGOR Application – (PREDICTION)

### VIGOR - TRAINING

- 5 Training Page
- 6 Choose Audio File
- 7 Speaker Diarization
- 8 Train the Target Speaker
- 9 Adding Target Speaker to Directory
- 10 Successfully added Speaker Message
- 11 Add to Present Speaker
- 12 Adding the New Speaker to the Old Speaker
- 13 Available Speakers
- 14 Previous Records
- 15 Enable Queue
- 16 Audio Queue

### VIGOR - PREDICTION

- 17 Prediction Home Page
- 18 Available Speakers - Prediction
- 19 New Calls
- 20 Old Calls
- 21 Automation Report

### NOTIFICATION

- 22 Notification App
- 23 Pop-up Window
- 24 Report Page

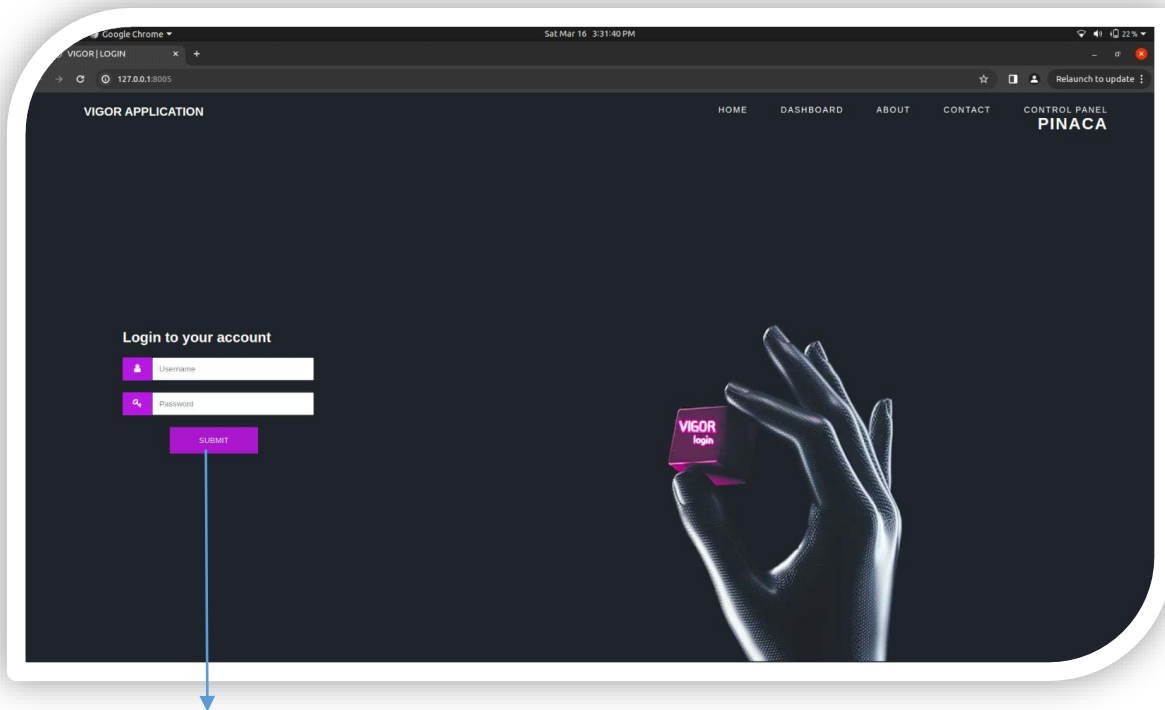
# VIGOR

Voice Based Identification and Generalised Object Recognition.

- Voice-based identification involves authenticating individuals by analyzing their unique voice patterns.
- Generalized object recognition, on the other hand, focuses on developing systems that can identify a wide range of objects across various scenarios without requiring specific training for each object.

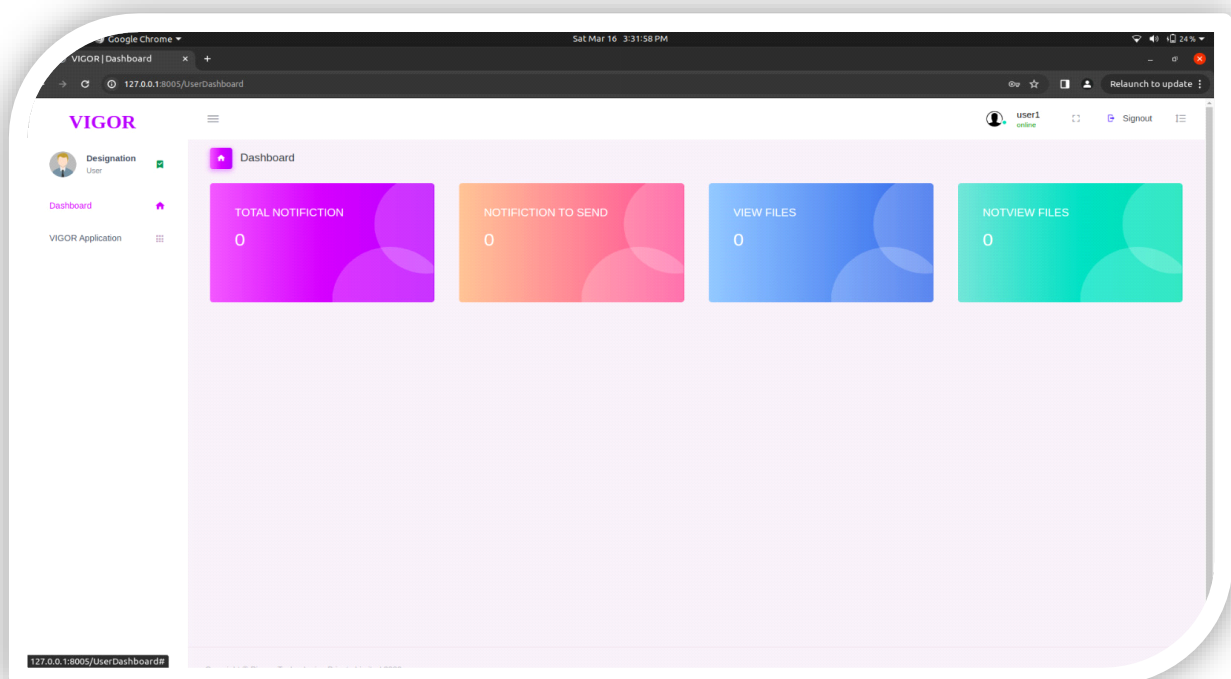
Both technologies play important roles in various applications, including security, automation, and user authentication.

# VIGOR LOGIN PAGE



"This is the login page for the Vigor application. Users are required to input their email and password. Once the details are entered, they should click on the 'Submit' button to proceed."

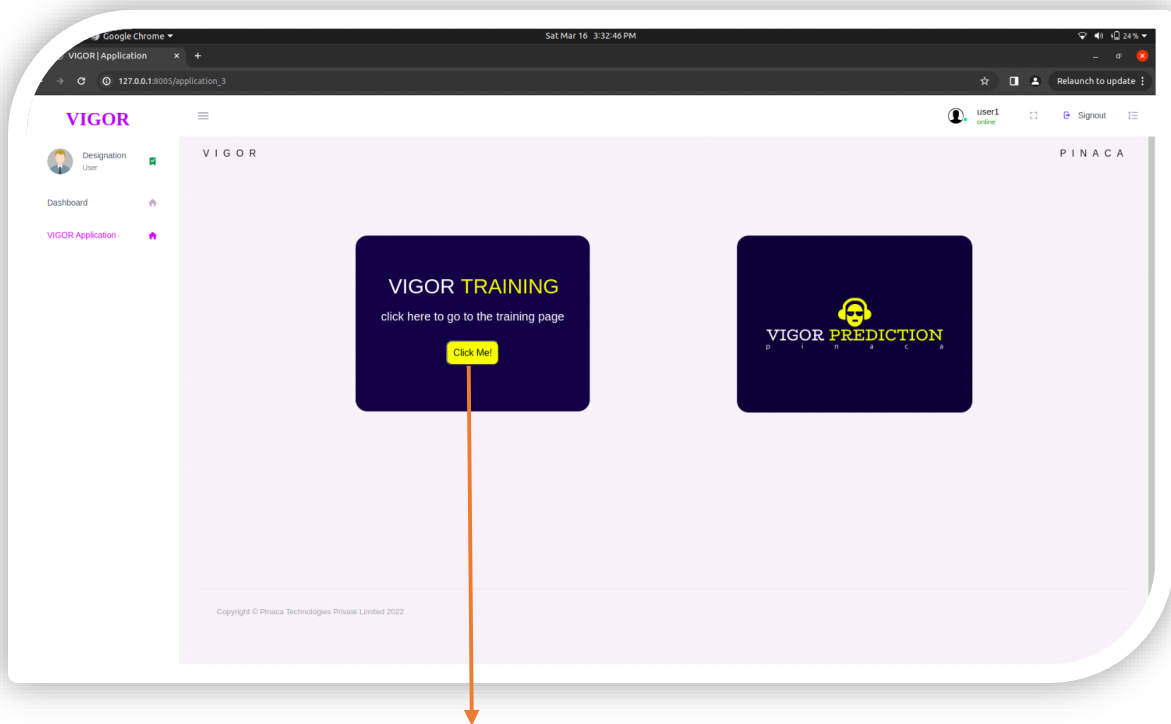
# VIGOR DASHBOARD



In the interface, there are four cards presenting the following information:

1. The first card provides an overview of the total number of notifications received by the user.
2. The second card displays the expected or pending notifications that the user is yet to receive.
3. The third card indicates the count of notifications that the user has already viewed.
4. The fourth card shows the number of notifications that remain unviewed by the user.

## VIGOR APPLICATION – (TRAINING)



"After selecting the VIGOR Application, you will encounter two cards:

1. The '**VIGOR Training**' card: Clicking on '**click me**' redirects you to the training page.
2. The '**VIGOR Prediction**' card: Clicking on '**click me**' redirects you to the prediction page."

Here, the initial selection is made by choosing the training card.

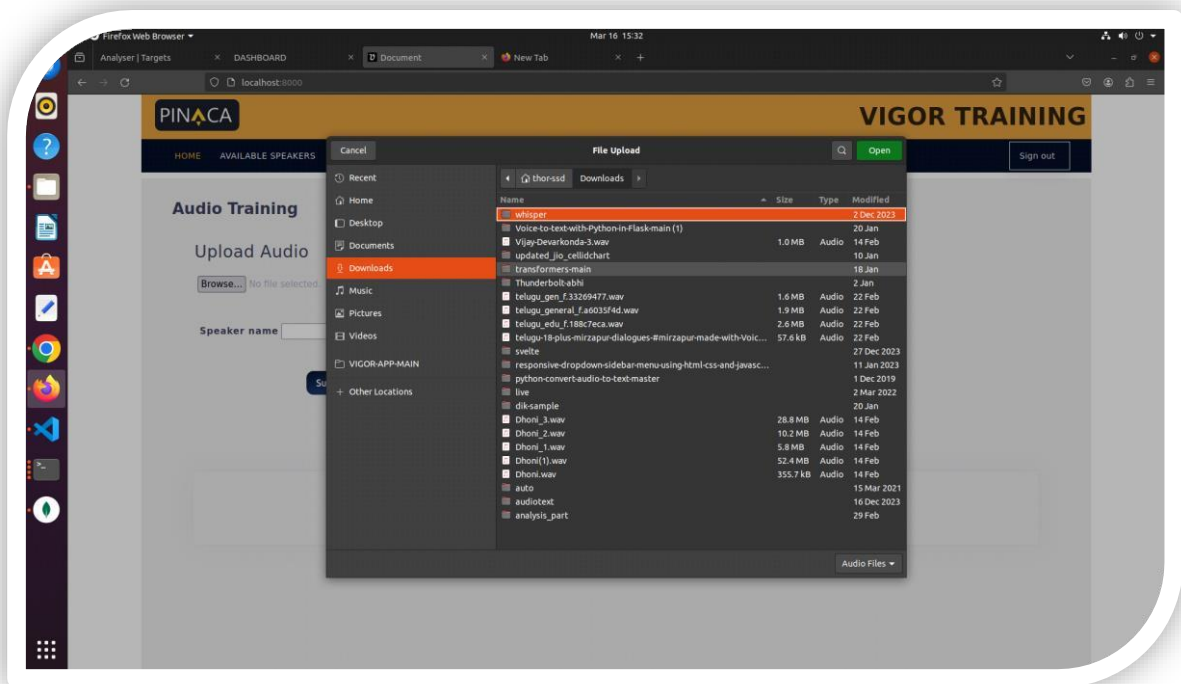
# TRAINING PAGE

The screenshot displays the VIGOR TRAINING web interface. At the top, there is a navigation bar with a logo on the left and the text 'VIGOR TRAINING' on the right. Below the navigation bar, there are two main sections: 'Audio Training' and 'Merge Audio Files'. The 'Audio Training' section includes an 'Upload Audio' button, an 'Enable Queue' toggle switch, a 'Browse...' button, and input fields for 'Speaker name' and 'Target number'. A 'Submit' button is located at the bottom of this section. The 'Merge Audio Files' section includes a 'Choose Audio File(s)' button, a 'Browse...' button, an 'Output File Name' input field, and a 'Merge' button. At the bottom of the page, there is a copyright notice: 'Copyright © 2022 Pinaca Technologies Private Limited , All rights reserved.'

On the training page, you have the following options:

- 1. Individual Training:** Upload an audio file to train a specific person (referred to as X). After uploading, enter the nick name and phone number associated with that audio file, then click the submit button.
- 2. Batch Training with Queue Enable:** Utilize the "Enable Queue" feature to train multiple audio files simultaneously by queuing them up for processing.
- 3. Merge Functionality:** Merge two or more audio files into a single file using the "Merge" option. Specify the output file name in the designated "Output File Name" field.

# CHOOSE AUDIO FILE

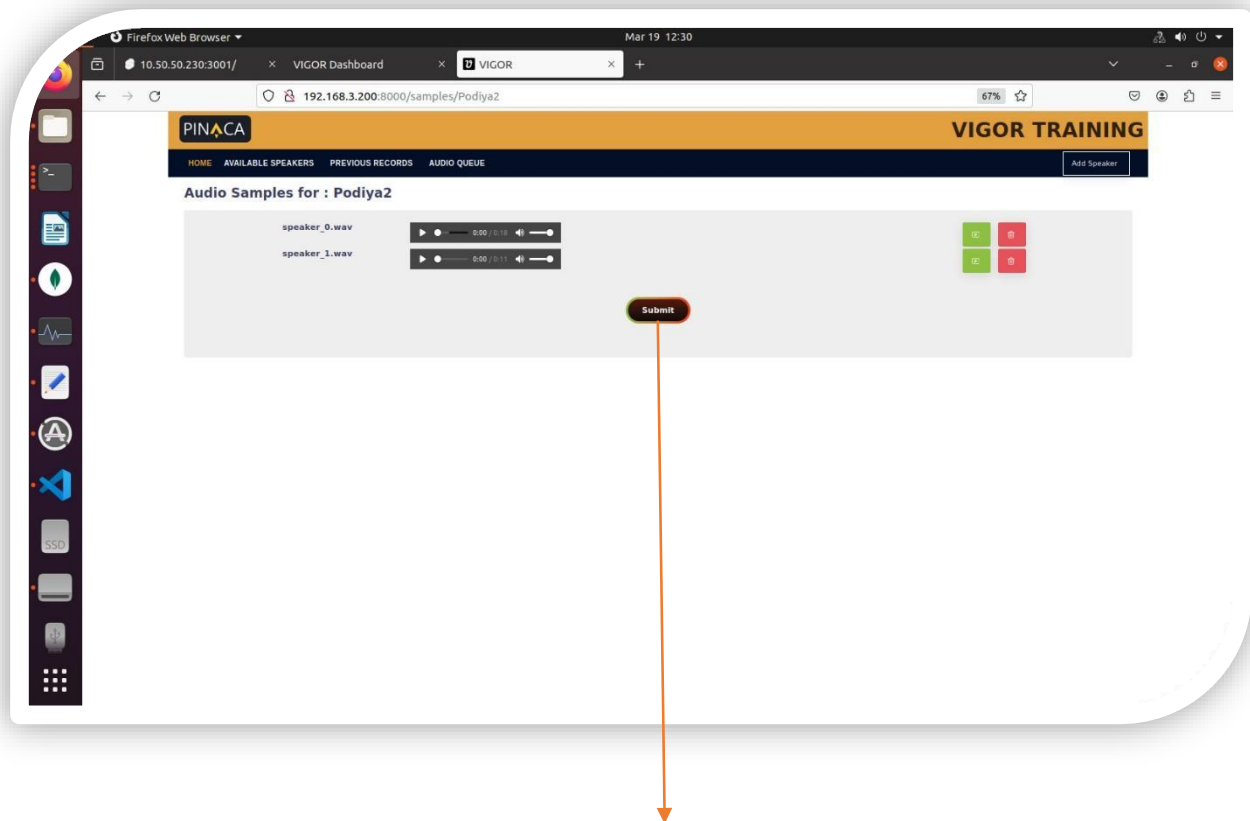


To choose an audio file, follow these steps:

1. Look for a button or an area labeled "**Choose File**"
2. Click on this button or area to open a file selection dialog box.
3. Navigate to the location on your device where the audio file is stored.
4. Select the desired audio file from the dialog box by clicking on it.
5. After selecting the file, click on the "Open" button in the dialog box to confirm your selection.

Once you've completed these steps, the audio file will be chosen and ready for further actions such as training, processing, or uploading.

# SPEAKER DIARIZATION



After submitting the audio file, the system will automatically separate all speakers present in the given audio file.

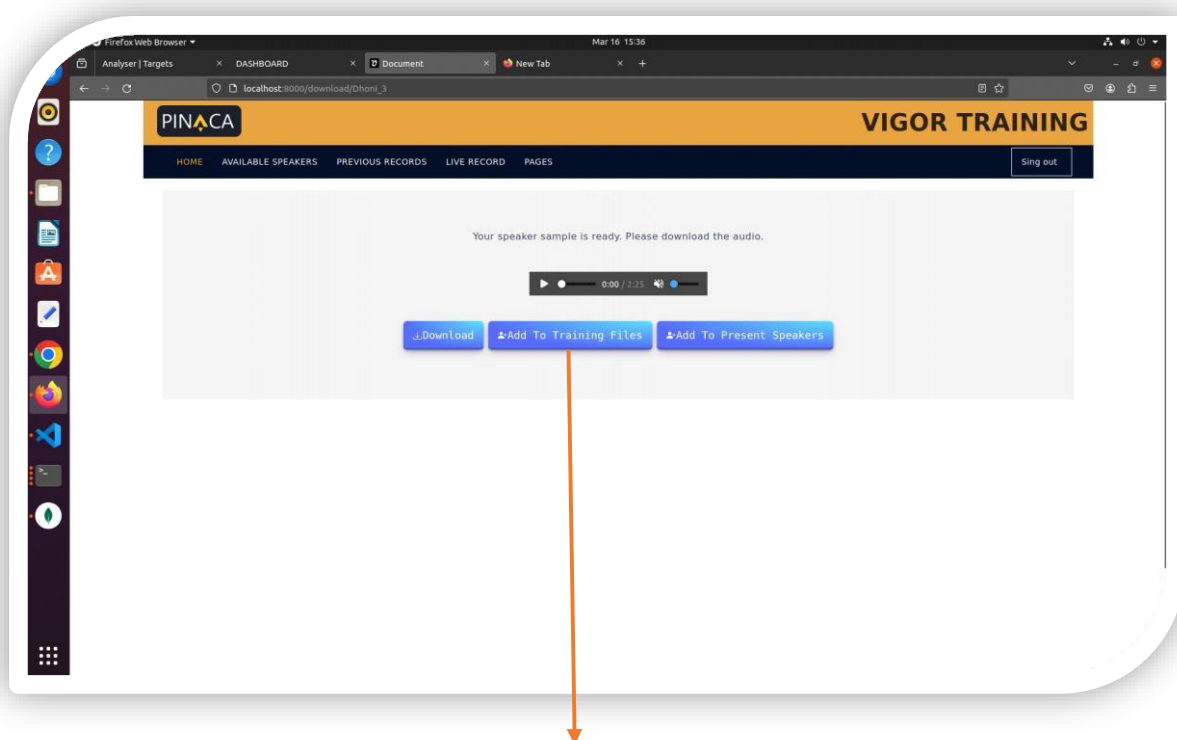
For example:

- If there is an audio file where speakers X and Y are engaged in a conversation, the system will divide the audio into two separate files, one for speaker X and another for speaker Y.

Once the separation is done, you can listen to each audio file and extract the target speaker's content. After extracting the desired speaker's content, you can delete the remaining portions from each file. Finally, click the submit button to complete the process.



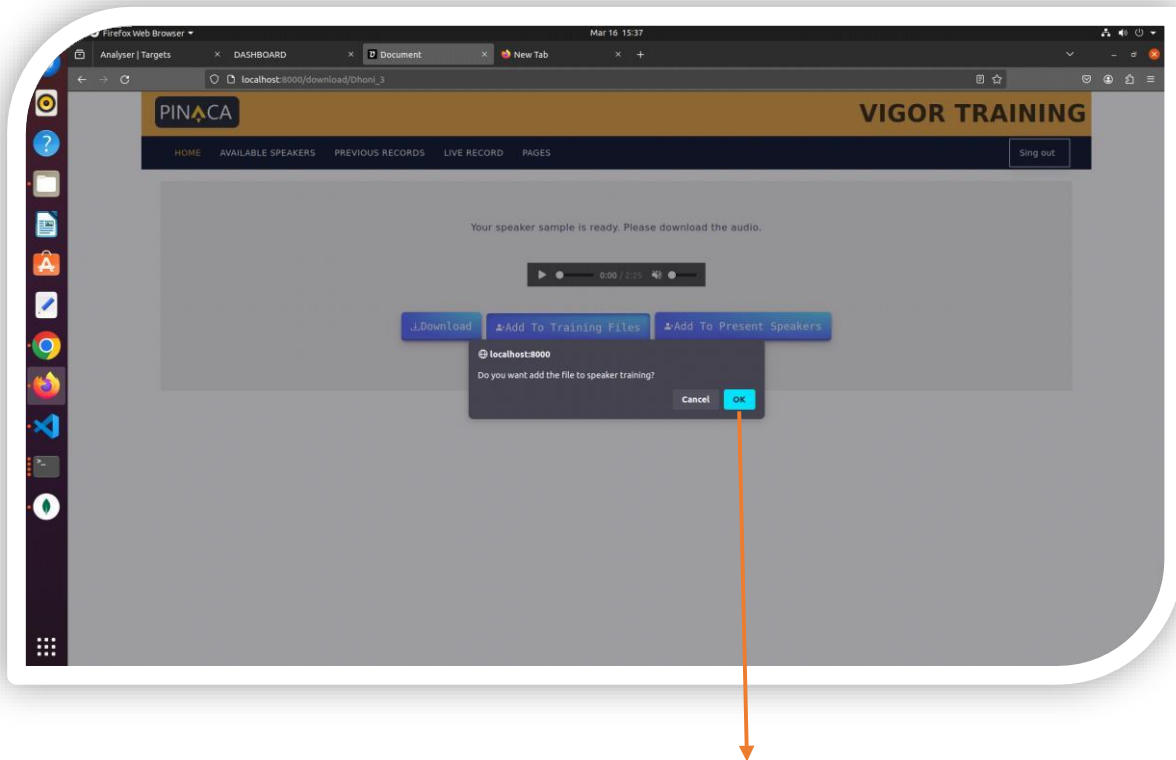
# TRAIN THE TARGET SPEAKER



Here are the options available after identifying the target speaker:

1. **Download Button:** Allows you to download the audio file of the target speaker.
2. **Add to Train Button:** Enables training of the target speaker's voice into our model, enhancing its ability to recognize and predict audio files matching the target speaker's voice.
3. **Add to Previous Speaker Button:** Provides the option to include the audio file of the target speaker to the existing speaker's voice in our training dataset.

Now, we should proceed by clicking the "**Add to Train**" button to train the target speaker into our model.



After clicking the "**Add to Train**" button, a popup message will appear with the following instruction:

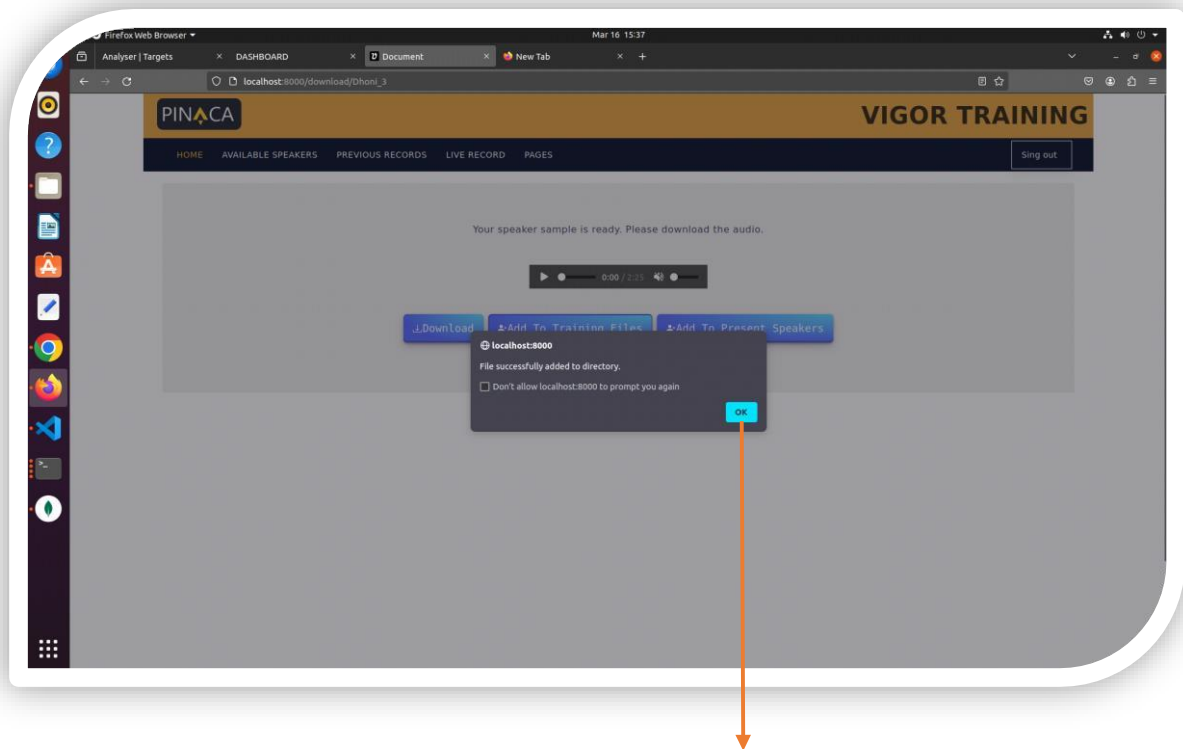
**"Do you want to add the file to speaker training?"**

You will have two options in this instruction:

1. **OK Button:** Clicking this button will proceed with the further process.
2. **Cancel Button:** Clicking this button will stop the process and return to the previous state.

Please click the "**OK**" button to continue with the training process.

# ADDING TARGET SPEAKER TO DIRECTORY

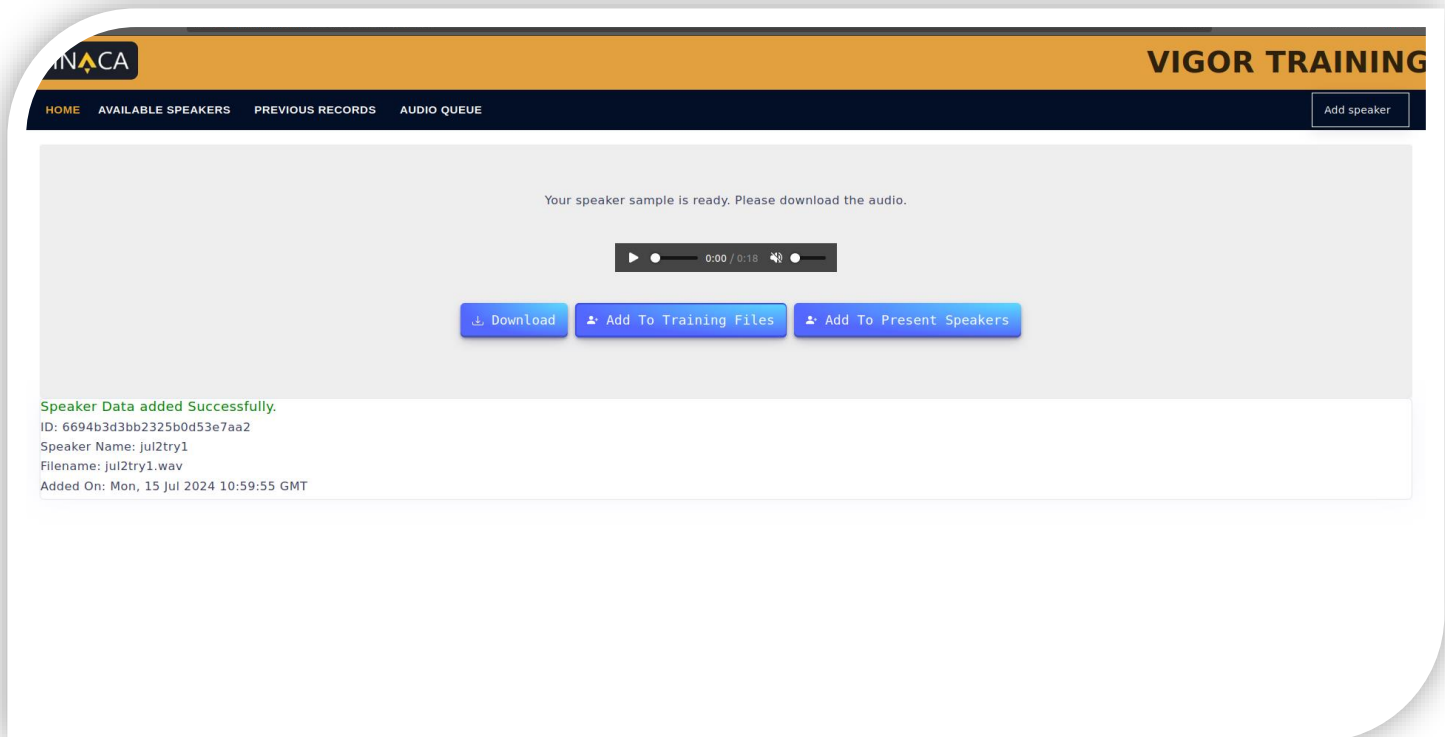


After clicking "**OK**" in the previous popup, another message will appear stating:

**"File successfully added to directory."**

You will find an "**OK**" button in this message. Please click the "**OK**" button to acknowledge the successful addition of the file to the directory.

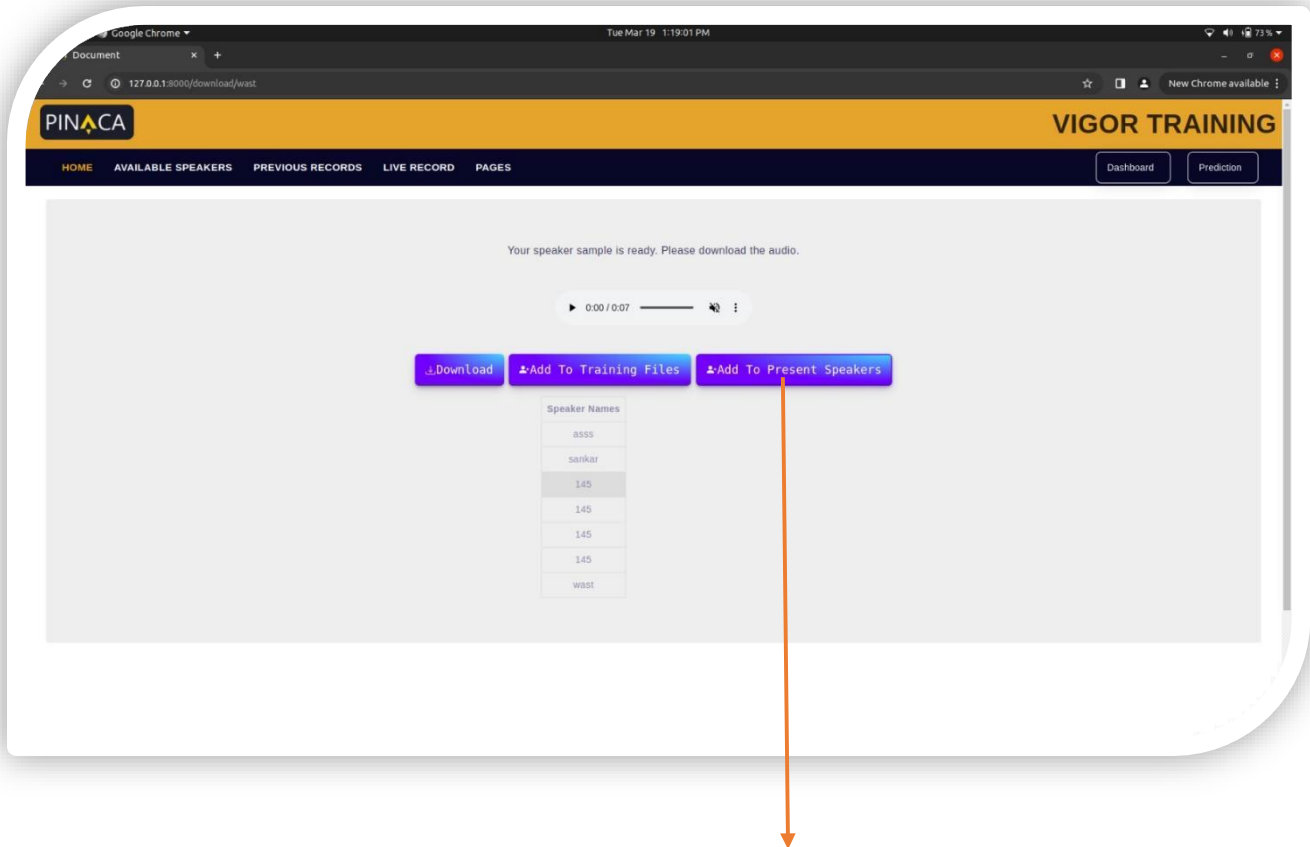
# SUCCESSFULLY ADDED SPEAKER MESSAGE



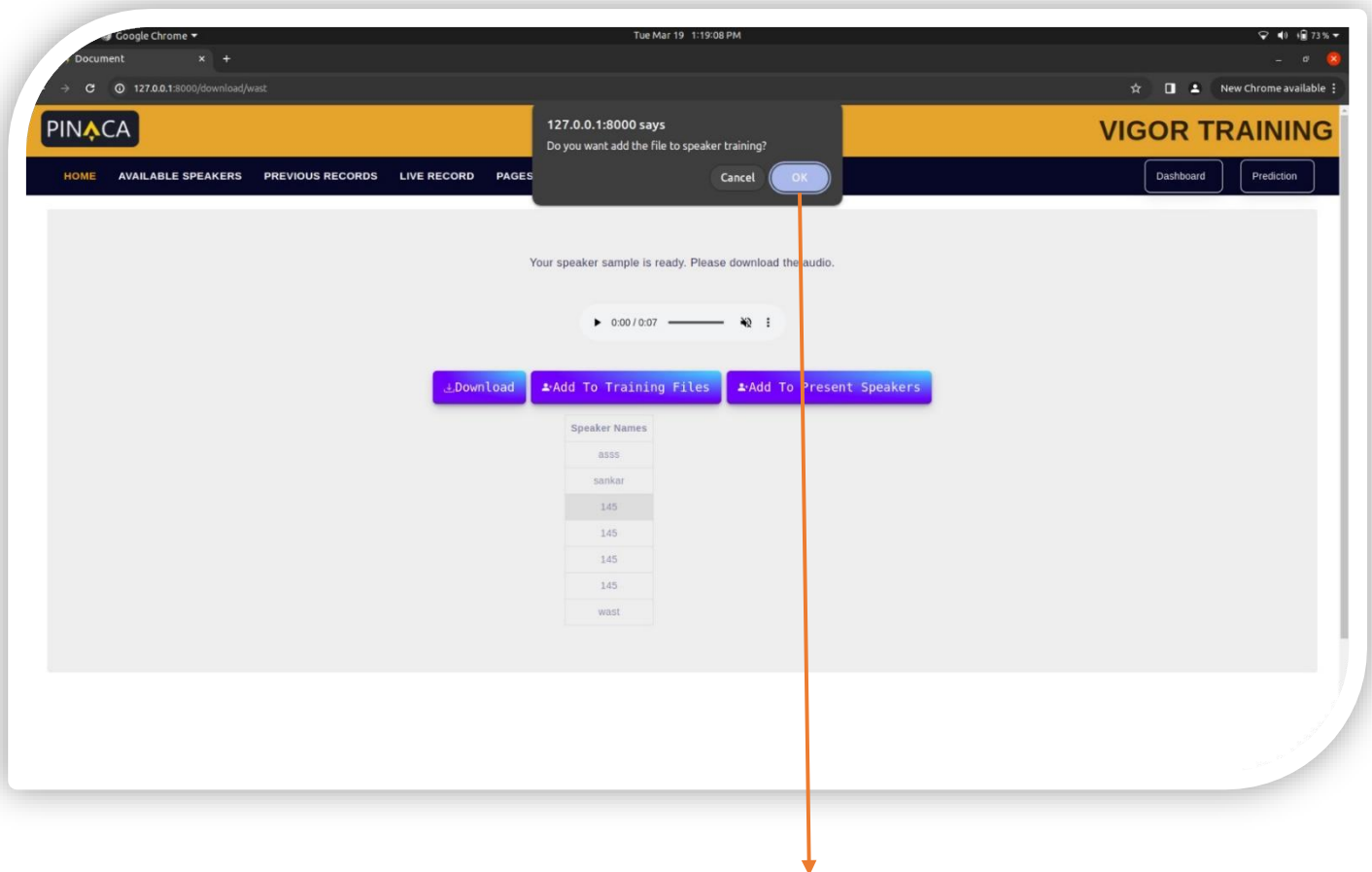
After clicking "**OK**" on the previous popup confirming "**File successfully added to directory.**" you will receive a message stating that the Speaker Data has been added successfully.

1. **Speaker Data Added Successfully:** The message confirms that the speaker data has been added successfully.
2. **Speaker Name:** Displays the name of the speaker provided during training on the homepage.
3. **Upload Date and Time:** Specifies the date and time when the audio file was uploaded.
4. **Audio File Name:** Provides the name of the uploaded audio file.

# ADD TO PRESENT SPEAKER



When you click on the **"Add to Previous Speaker"** button, it will present a list of trained speaker names. Then after you should select the specific speaker name from this list to add the trained speaker to the previously stored speaker's data.



After selecting a speaker name from the displayed list, you will receive the following instruction:

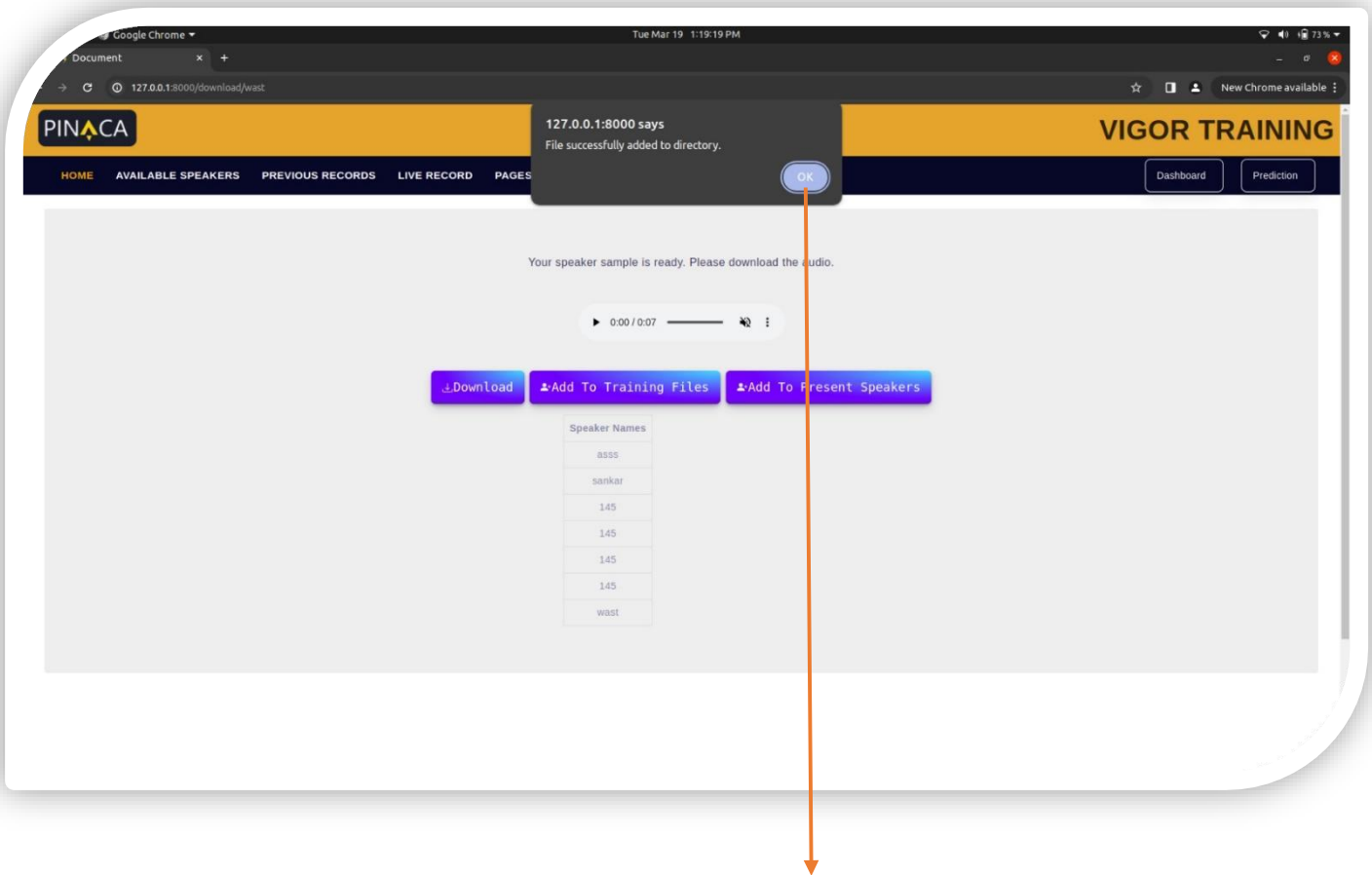
**"Do you want to add the file to speaker training?"**

You will have two options in this instruction:

1. **OK Button:** Clicking this button will proceed with the further process.
2. **Cancel Button:** Clicking this button will stop the process and return to the previous state.

Please click the **"OK"** button to continue with the training process.

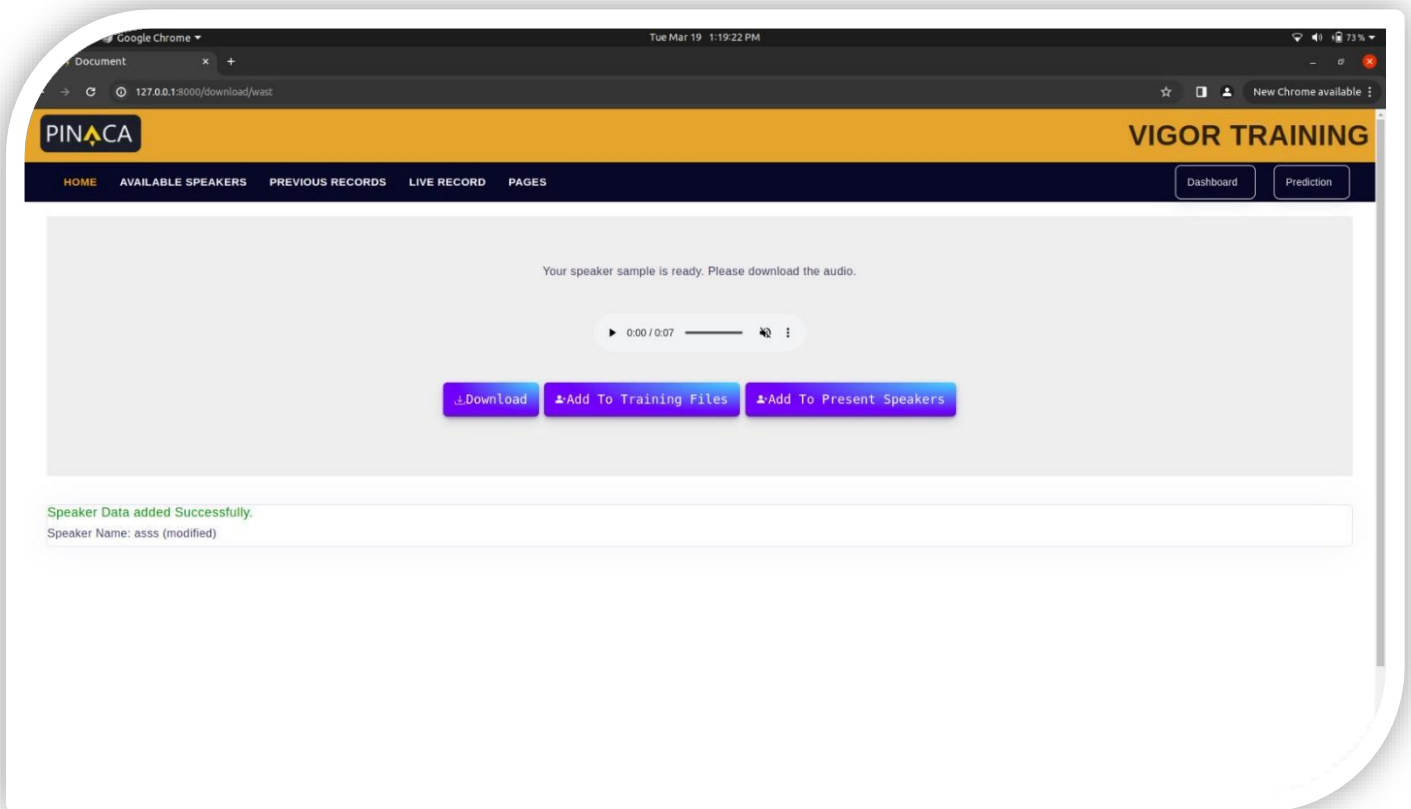
# ADDING THE NEW SPEAKER TO OLD SPEAKER



After clicking "**OK**" in the previous popup, another message will appear stating:

**"File successfully added to directory."**

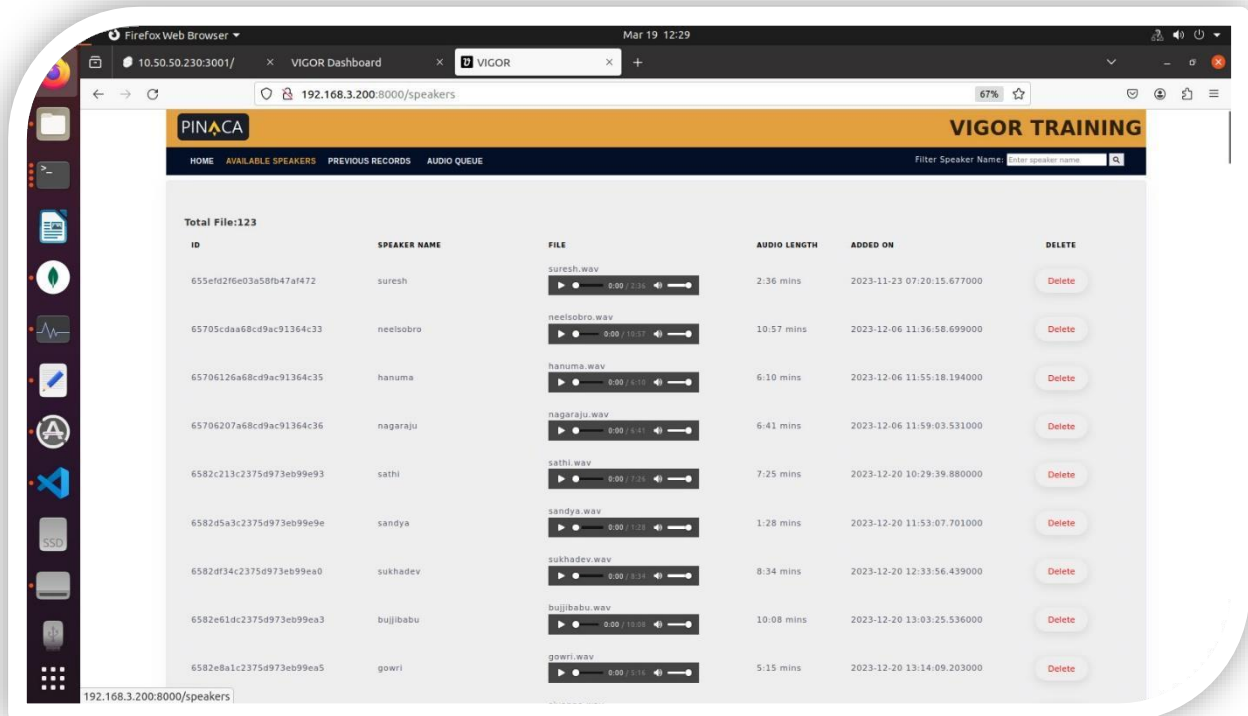
You will find an "**OK**" button in this message. Please click the "**OK**" button to acknowledge the successful addition of the file to the directory.



After clicking “**OK**” on the previous popup confirming “**File successfully added to directory.**” You will receive a message stating that the speaker data has been added successfully.



# AVAILABLE SPEAKERS



Here we can see the trained target speaker's audio files.

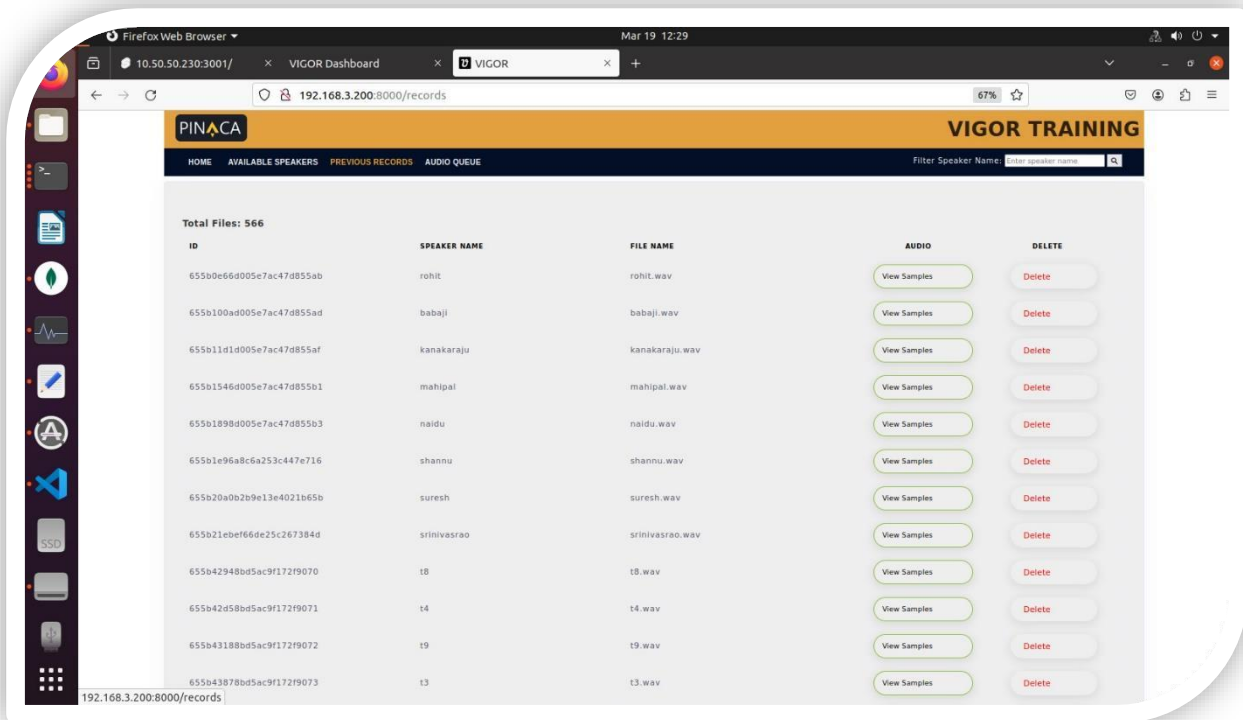
## 1. Available Speakers:

- Displays all trained target speakers.
- Allows listening to the trained target speaker's audio.
- Provides the specific speaker's name for identification.
- Shows the audio length of the target speaker's recordings.

## 2. Delete Option:

- Enables the removal of trained target speakers from the account if needed.

# PREVIOUS RECORDS



In the "**Previous Records**" or we can also call "History" section, you can access past audio file records. Here are the details:-

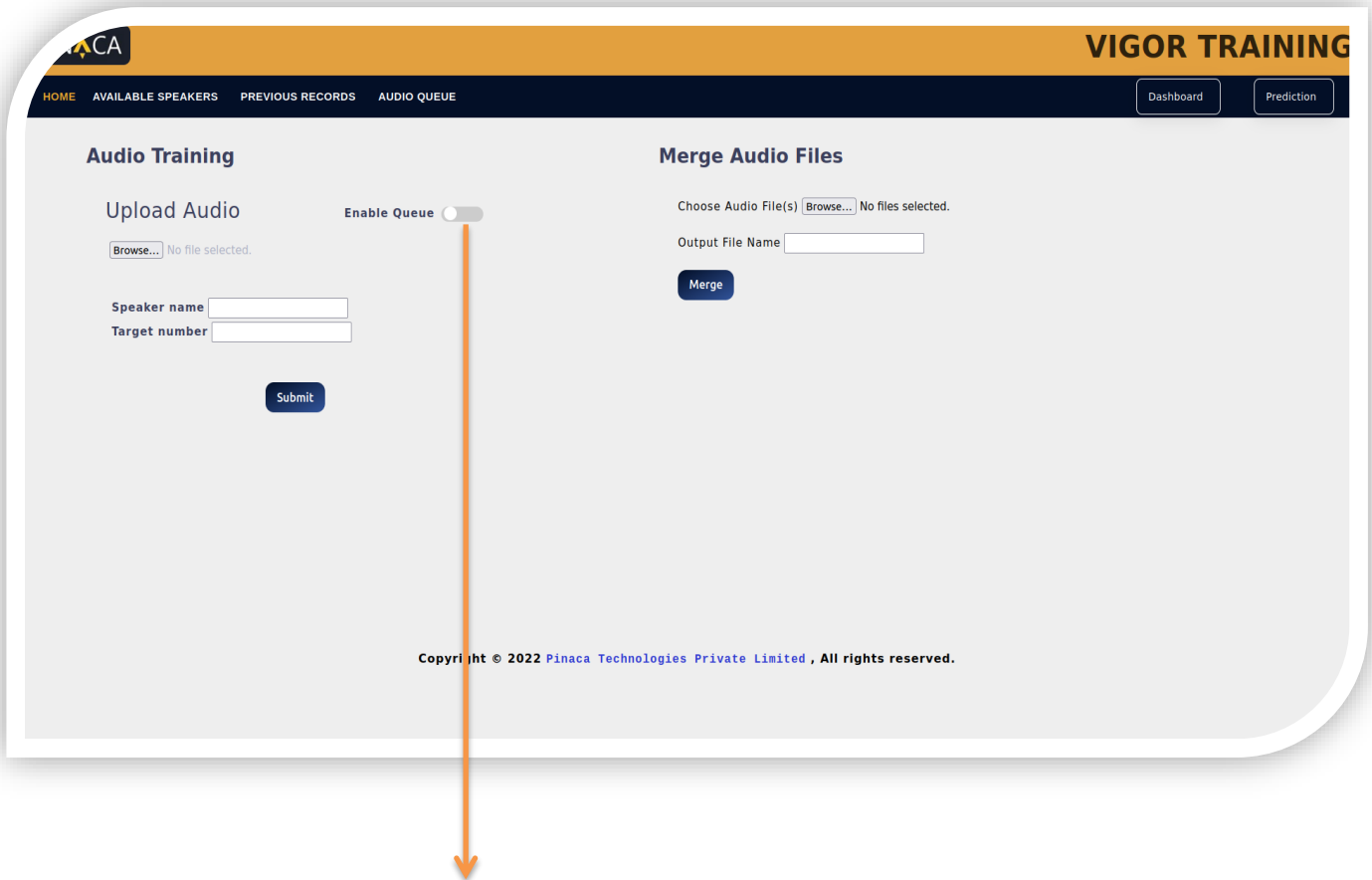
## 1. Record Storage:

- All records of given audio files are stored here.
- Each record includes details such as speaker information and audio content.

## 2. Functionalities:

- Delete Option: Allows deletion of specific audio files from the record.
- View Samples: Redirects to a **speaker diarization** page where speakers are individually identified and segmented.

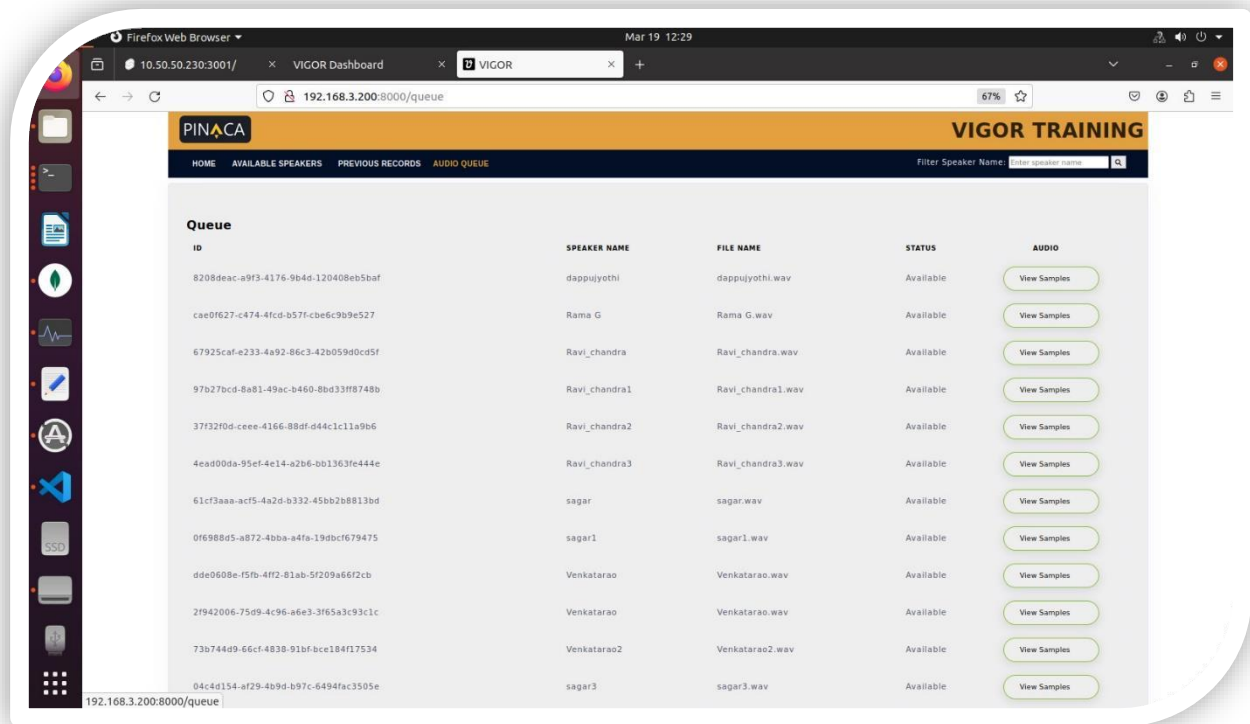
# ENABLE QUEUE



The "**Enable Queue**" feature allows training multiple files efficiently by queuing them sequentially.

When the "Enable Queue" toggle button is activated, the system processes and queues all selected files for training. This feature operates independently of the "Previous Records" section, ensuring that queued files are managed separately for streamlined training of multiple files in succession. And this files are showing in the **AUDIO QUEUE**.

# AUDIO QUEUE



In the "**Audio Queue**" section, you can view the audio files that have been added to the enable queue for processing.

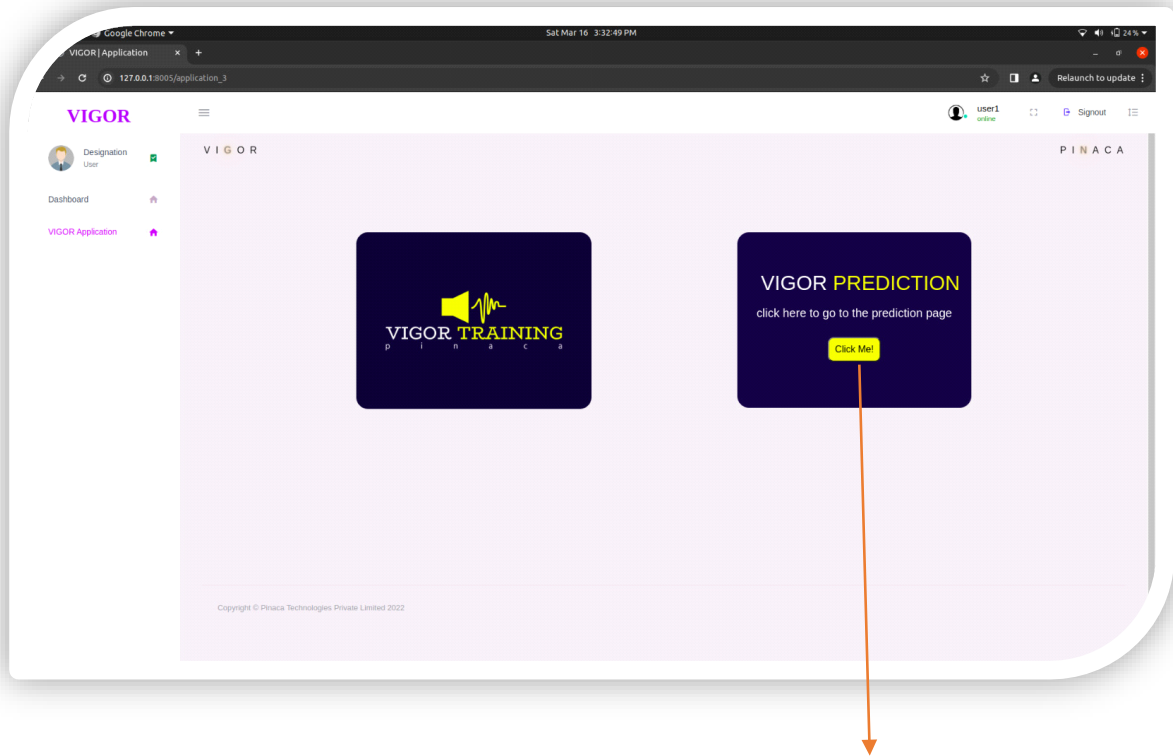
## 1. Audio Queue Display:

- Shows the audio files added to the enable queue for processing.

## 2. Functionality of "**View Sample**" Button:

- Clicking the "**View Sample**" button displays the individual speakers' segments from the selected audio file.
- This action redirects to the "**Speaker Diarization**" page, providing a segmented view of speakers within the audio file.

# VIGOR APPLICATION – (PREDICTION)

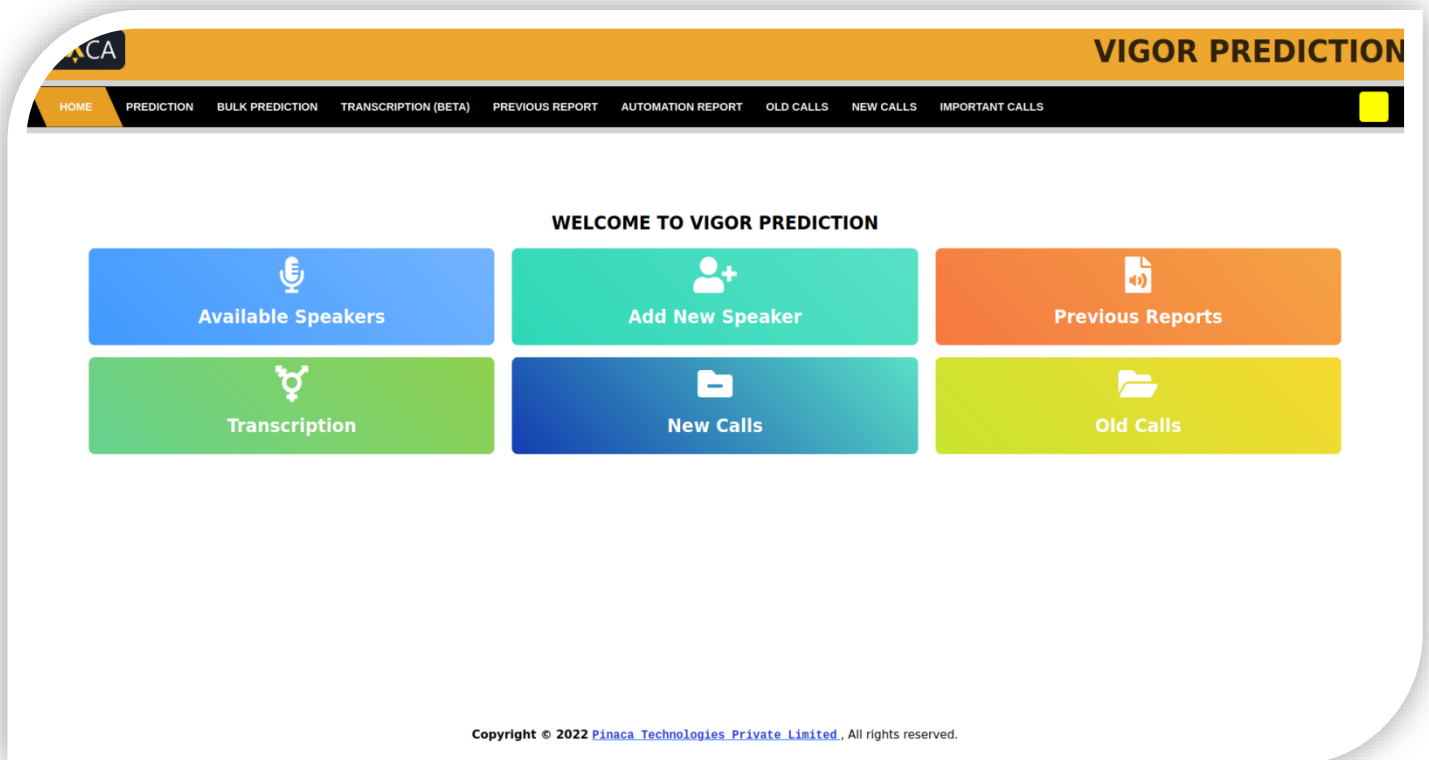


"After selecting the VIGOR Application, you will encounter two cards:

1. The '**VIGOR Training**' card: Clicking on '**click me**' redirects you to the training page.
2. The '**VIGOR Prediction**' card: Clicking on '**click me**' redirects you to the prediction page."

Here, the initial selection is made by choosing the Prediction card.

# PREDICTION HOME PAGE



On the homepage, you will find six options available:

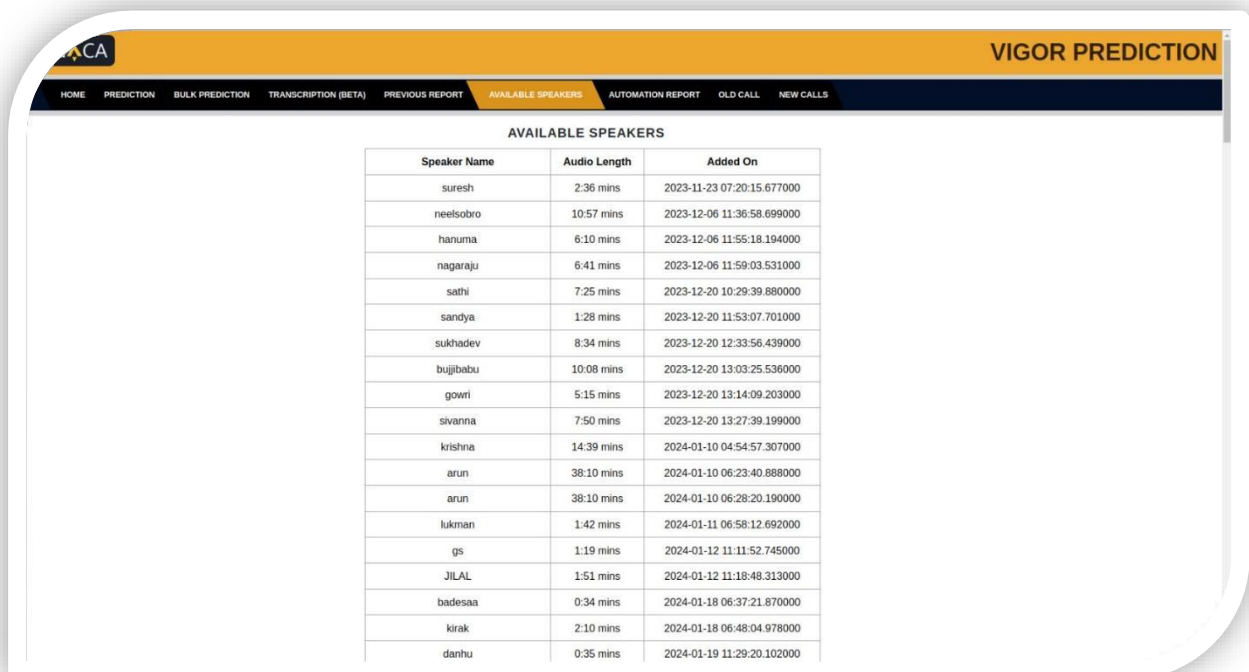
1. **Available Speakers:** This option redirects you to a page displaying speakers whose voices have been used for training in the VIGOR Training section.
2. **Add New Speaker:** Clicking this option will take you to the VIGOR Training homepage where you can add a new speaker for training.
3. **Previous Reports:** This option provides access to the history of audio files submitted for prediction in the system.
4. **Transcription:** Redirects to the Transcription page, where you can upload a file and receive the English translation of the audio.

5. **New Calls:** Redirects to the New Calls page, where you can view recent audio files.

6. **Old Calls:** Redirects to the Old Calls page, where you can view previously reviewed audio files.

So now select the **Available Speakers**.

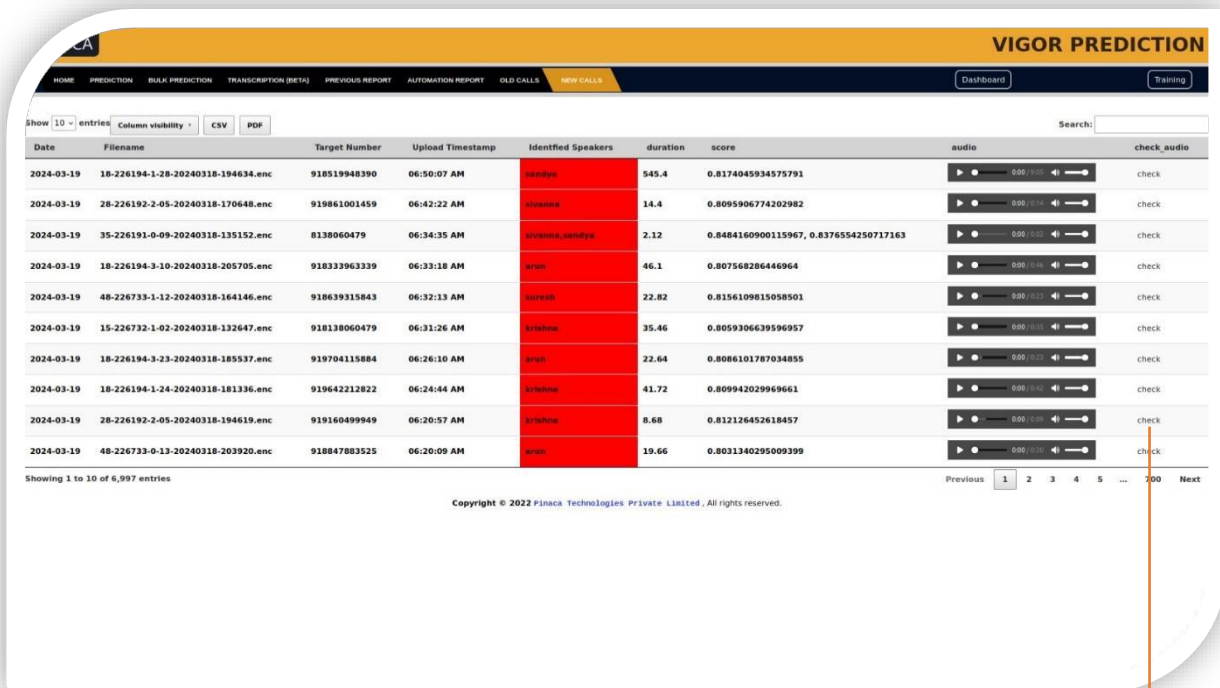
## AVAILABLE SPEAKERS – (PREDICTION)



AVAILABLE SPEAKERS		
Speaker Name	Audio Length	Added On
suresh	2:36 mins	2023-11-23 07:20:15.677000
neelsobro	10:57 mins	2023-12-06 11:36:58.699000
hanuma	6:10 mins	2023-12-06 11:55:18.194000
nagaraju	6:41 mins	2023-12-06 11:59:03.531000
sathi	7:25 mins	2023-12-20 10:29:39.880000
sandya	1:28 mins	2023-12-20 11:53:07.701000
sukhadev	8:34 mins	2023-12-20 12:33:56.439000
bujibabu	10:08 mins	2023-12-20 13:03:25.536000
gowri	5:15 mins	2023-12-20 13:14:09.203000
sivanna	7:50 mins	2023-12-20 13:27:39.199000
krishna	14:39 mins	2024-01-10 04:54:57.307000
arun	38:10 mins	2024-01-10 06:23:40.888000
arun	38:10 mins	2024-01-10 06:28:20.190000
lukman	1:42 mins	2024-01-11 06:58:12.692000
gs	1:19 mins	2024-01-12 11:11:52.745000
JILAL	1:51 mins	2024-01-12 11:18:48.313000
badesaa	0:34 mins	2024-01-18 06:37:21.870000
kirak	2:10 mins	2024-01-18 06:48:04.978000
danhu	0:35 mins	2024-01-19 11:29:20.102000

In the "Available Speakers" section, you can view the speakers who have been trained in the VIGOR Training module, along with the duration of their respective trained audio files.

# NEW CALLS – (PREDICTION)



Date	Filename	Target Number	Upload Timestamp	Identified Speakers	duration	score	audio	check_audio
2024-03-19	18-226194-1-28-20240318-194634.enc	918519948390	06:50:07 AM	crash	545.4	0.8174045934575791		check
2024-03-19	28-226192-2-05-20240318-170648.enc	919861001459	06:42:22 AM	crash	14.4	0.8095906774202982		check
2024-03-19	35-226191-0-09-20240318-135152.enc	8138060479	06:34:35 AM	crash, sandy	2.12	0.8484160900115967, 0.8376554250717163		check
2024-03-19	18-226194-3-10-20240318-205705.enc	918333963339	06:33:18 AM	crash	46.1	0.807568286446964		check
2024-03-19	48-226733-1-12-20240318-164146.enc	918639315843	06:32:13 AM	crash	22.82	0.8156109815058501		check
2024-03-19	15-226732-1-02-20240318-132647.enc	918138060479	06:31:26 AM	crash	35.46	0.8059306439596957		check
2024-03-19	18-226194-3-23-20240318-185537.enc	919704115884	06:26:10 AM	crash	22.64	0.8086101787034855		check
2024-03-19	18-226194-1-24-20240318-181336.enc	919642212822	06:24:44 AM	crash	41.72	0.809942029969661		check
2024-03-19	28-226192-2-05-20240318-194619.enc	919160499949	06:20:57 AM	crash	8.68	0.812126452618457		check
2024-03-19	48-226733-0-13-20240318-203920.enc	918847883525	06:20:09 AM	crash	19.66	0.803134025009399		check

Showing 1 to 10 of 6,997 entries

Copyright © 2022 Pinaca Technologies Private Limited. All rights reserved.

Previous 1 2 3 4 5 ... 700 Next

On the "New Calls" page, you can access all predicted audio files that have been matched with trained speakers. Here are the details provided on this page:

1. **Audio File Name:** Displays the name of the particular audio file.
2. **Target Number:** Indicates the target person or speaker associated with the audio file.
3. **Uploaded Time:** Shows the timestamp of when the audio file was uploaded.
4. **Identified Speakers:** Lists the speakers identified within the predicted audio file.
5. **Duration:** Specifies the length of the audio file.
6. **Audio Playback:** Allows you to listen to the audio file.
7. **Check Button:** After listening to the audio file, click the check button to confirm that you have reviewed it. Upon clicking, the audio file will move to the "Old Calls" page.



# OLD CALLS – (PREDICTION)

Date	Filename	Target Number	Upload Timestamp	Identified Speakers	duration	score	audio
2024-02-29	48-226733-3-12-20240229-120849.enc	919542709306	06:42:28 PM	Sanjya	18.88	0.8054686856269836	
2024-02-29	18-226194-3-26-20240229-095251.enc	918977282983	03:36:05 PM	Arjuna	12.82	0.8126033991575241	
2024-02-29	28-226192-3-08-20240229-163655.enc	919160848888	01:41:13 PM	Ruthadav	25.86	0.8029978615897042	

Showing 1 to 3 of 3 entries

Copyright © 2022 Pinaca Technologies Private Limited, All rights reserved.

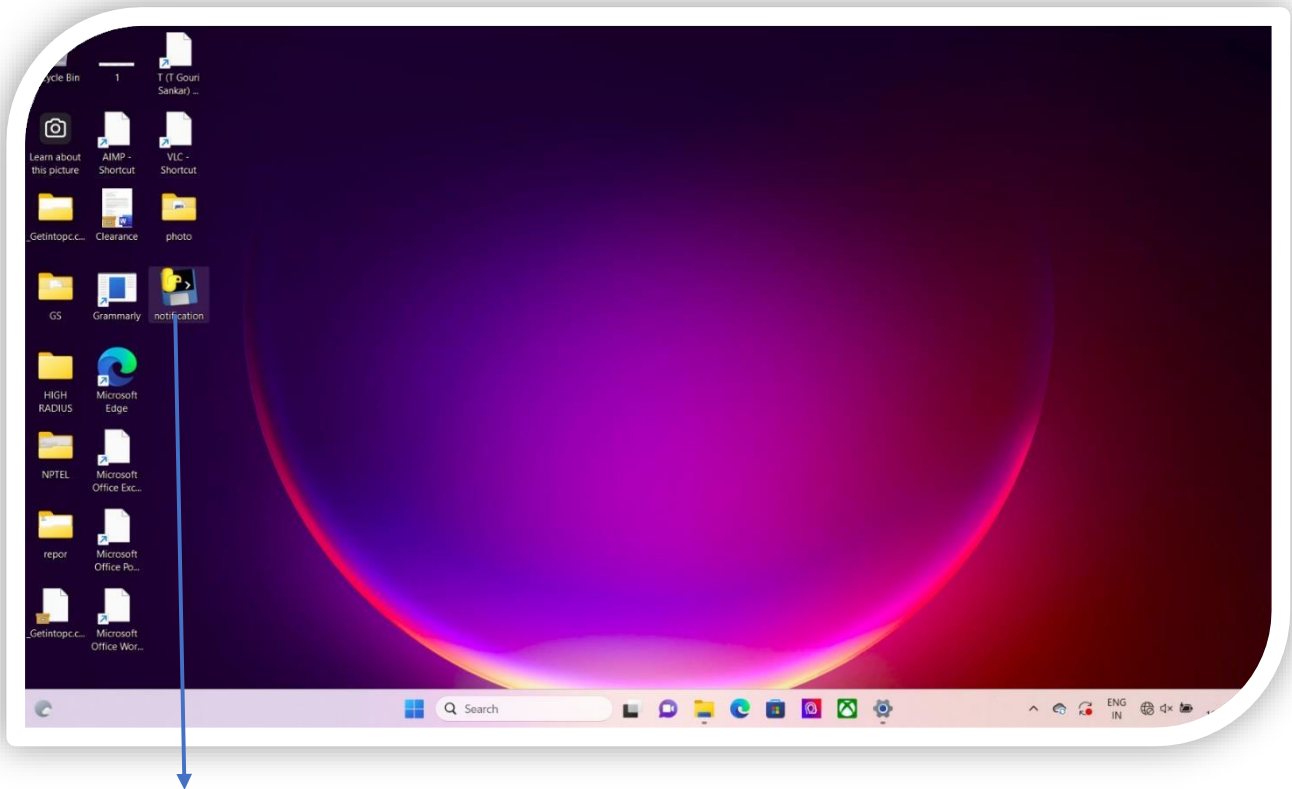
In the "Old Calls" page, you will find the audio files that you have already listened to in the "New Calls" page. These audio files are stored in the "Old Calls" section after being reviewed in the "New Calls" page.

# AUTOMATION REPORT – (PREDICTION)

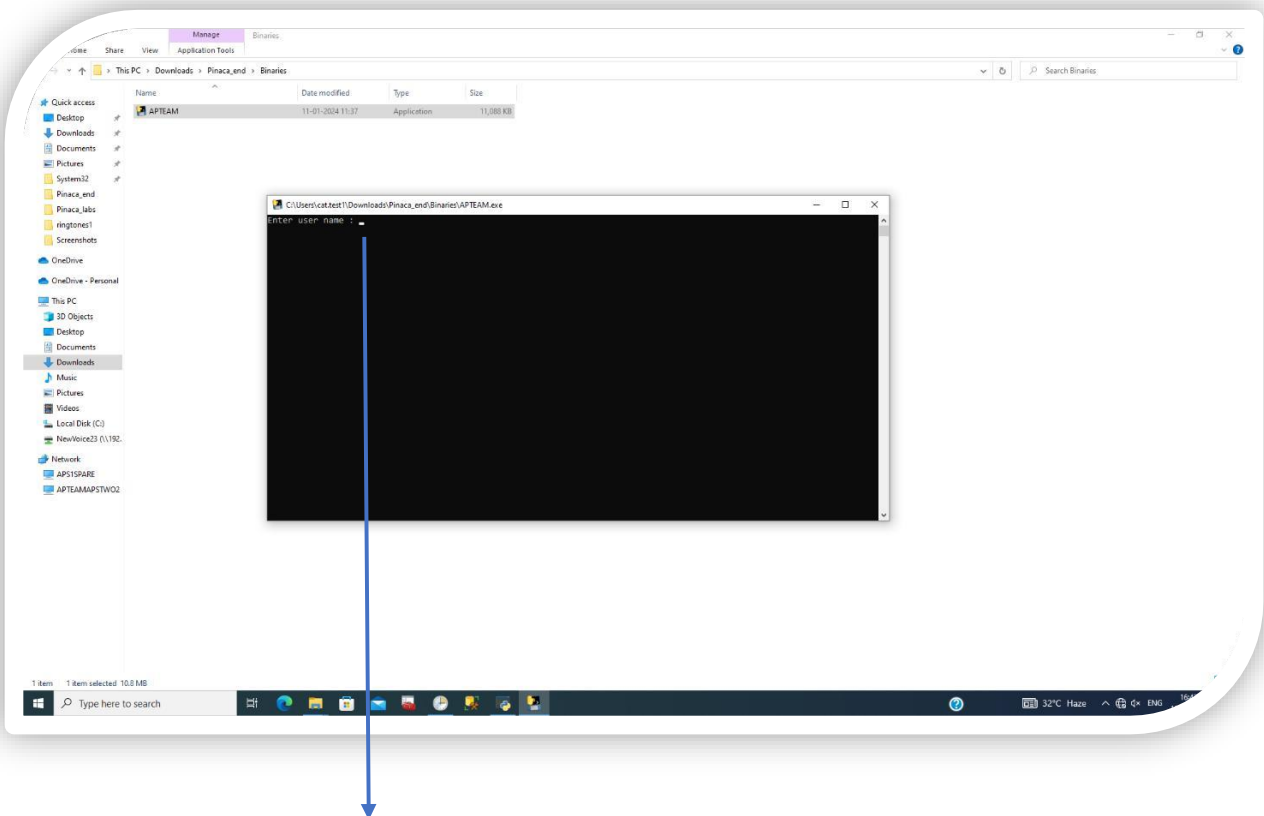
VIGOR PREDICTION								
HOME PREDICTION BULK PREDICTION TRANSCRIPTION (BETA) PREVIOUS REPORT AUTOMATION REPORT OLD CALLS NEW CALLS IMPORTANT CALLS								
how 10 entries Column visibility CSV PDF Search:								
Date	Filename	Target Number	Upload Timestamp	Identified Speakers	duration	score	audio	Transcription
2024-05-24	testaudio_client.wav		2024-05-24 13:00:01.034616	shannu,sankar,Babajl	1:04 min	0.7834632694721222, 0.7573407816834845, 0.7382322522490306		Transcription
2024-05-21	hindi_client_test.wav		2024-05-21 13:06:14.491189	shannu,Babajl	1:00 min	0.7471558354134225, 0.7204531196060531		Transcription
2024-03-28	audio.wav		2024-03-28 15:21:41.394085	wast	0:00 min	0.7252339124679565		Transcription
2024-03-22	14-226735-1-10-20230819-124137.wav		2024-03-22 12:16:58.426342	shannu	1:39 min	0.7066145848867007		Transcription
2024-03-22	wast.wav		2024-03-22 12:16:58.426342		0:33 min			Transcription
2024-03-22	14-226735-1-09-20230819-172602.wav		2024-03-22 12:16:58.426342		0:44 min			Transcription
2024-02-19	tyguhij.wav		2024-02-19 12:25:07.965514	asss,sankar,145	0:15 min	0.8673990744326083, 0.8665168886667831, 0.8447785062411808		Transcription
2024-02-19	wedsad.wav		2024-02-19 12:24:08.918151		0:31 min			Transcription
2024-02-13	wast1111.wav		2024-02-13 16:21:06.208943	Babajl,shannu	0:33 min	0.7070353261997245, 0.70369189891512		Transcription
						0.7542686686603302.		

The Automation Report offers a comprehensive overview of both previously reviewed audio files from the "Old Calls" section and newly predicted audio files from the "New Calls" section. Additionally, it features a transcription button that translates the audio into English when clicked.

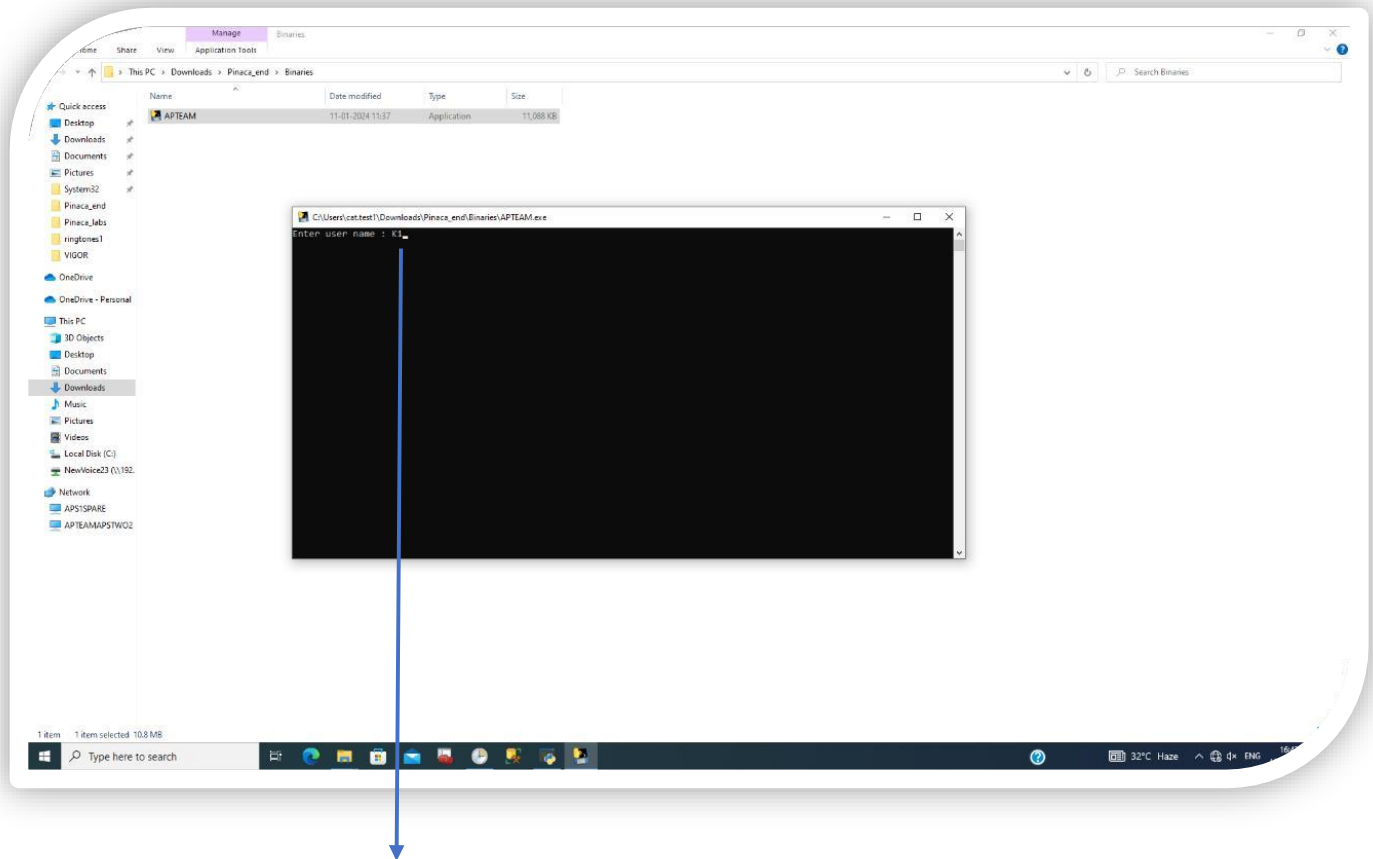
# NOTIFICATION APP



On the desktop, there is a **notification app** that we will utilize because it sends popup notifications instead of navigating to the VIGOR Prediction page. Double-click on it to launch the application.

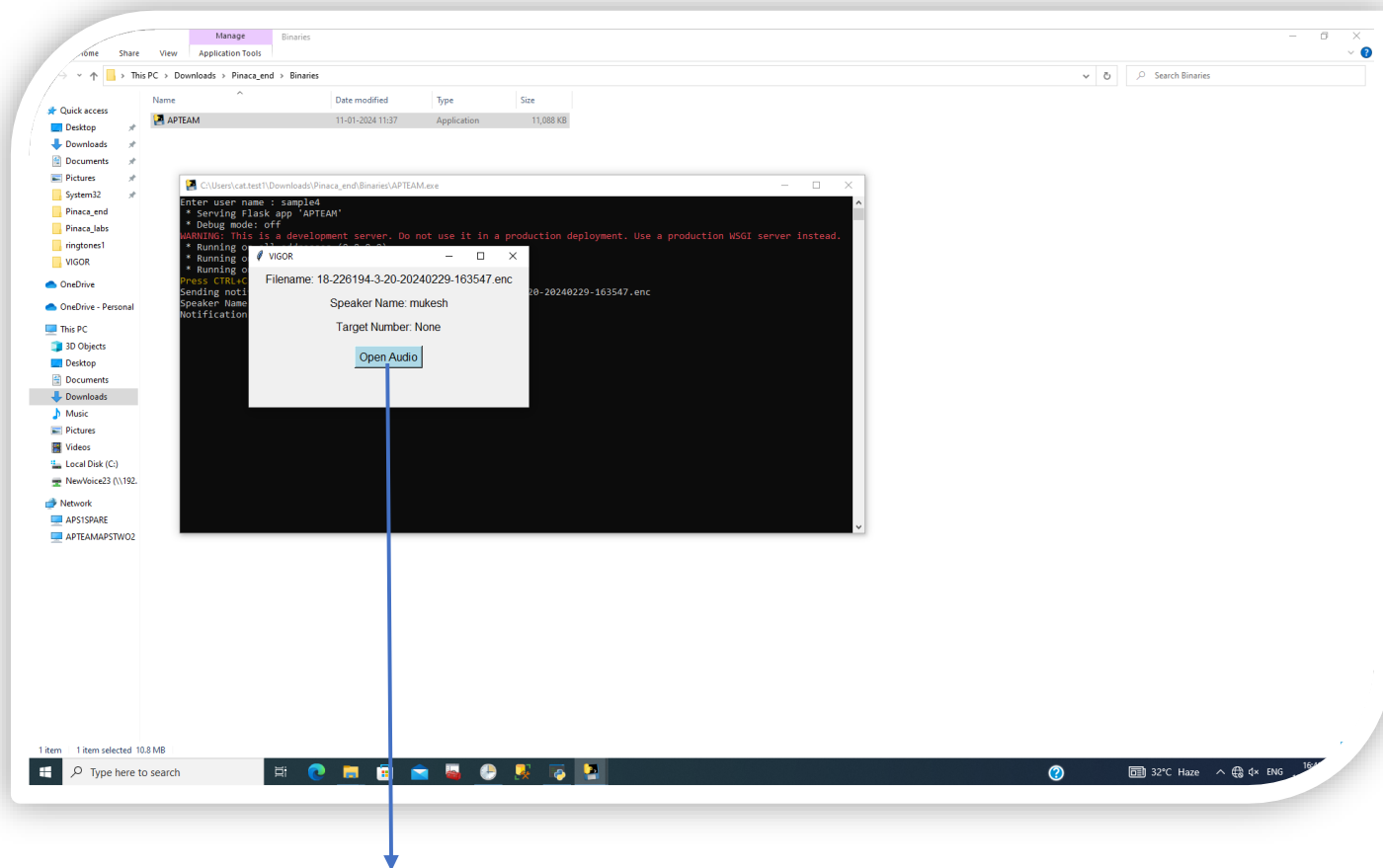


When the terminal prompts you with "**Enter the username:**", please ensure to enter your username in capital letters, for example, A2, A3, and so on.



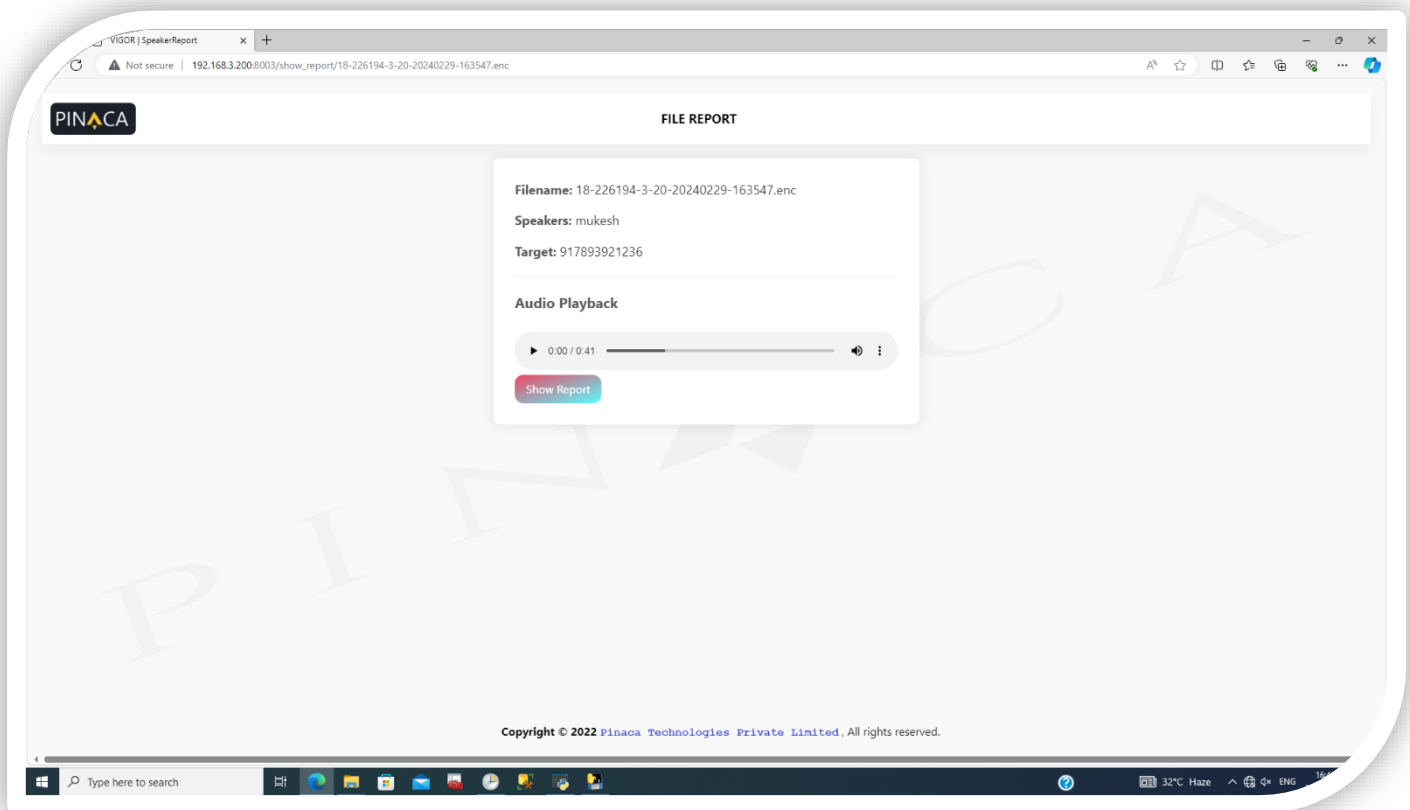
After entering your username, press the **Enter** key to execute the application.

# POP-UP WINDOW



After entering the username and pressing the "**Enter**" button on the keyboard, the application starts sending popup windows to the system. Each popup window displays the matched Speaker name, Target Number, and an "Open Audio" button. Clicking the "**Open Audio**" button redirects to the "Report Page" for further details and analysis.

# REPORT PAGE



Upon clicking the "**Open Audio**" button, you will be redirected to the file report page. Here, you can access information such as the matched Speaker name, Target number, and the Audio call. You have the option to listen to the audio call and verify the presence of the speaker in the call.