

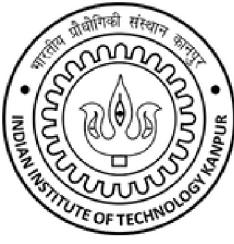


EE392A (UGP-II)

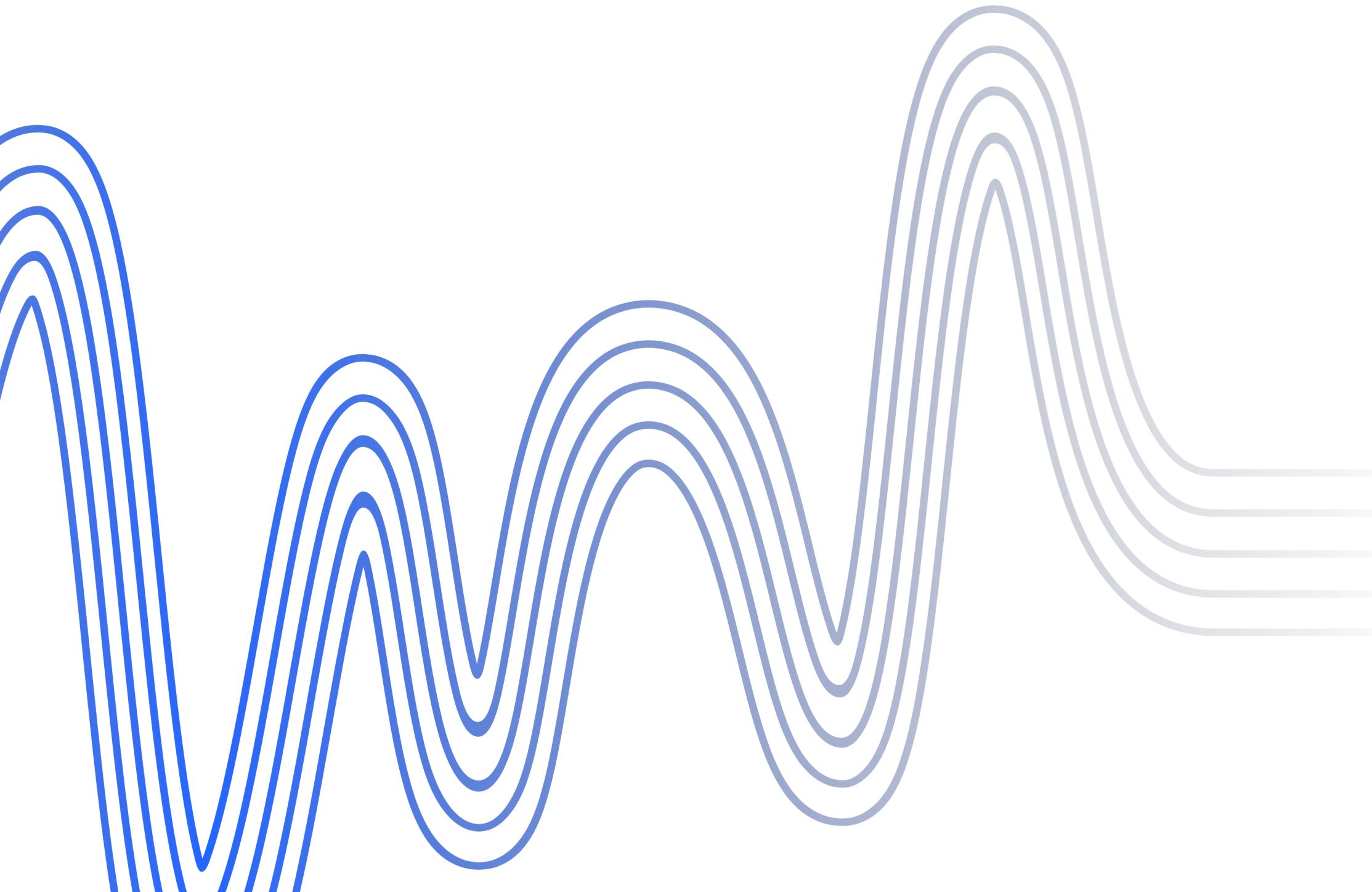
MELODY EXTRACTION & ANNOTATION TOOL

Under the Guidance of
Prof. Vipul Arora,
Associate Professor
Department of Electrical Engineering,
IIT Kanpur (U.P)

Submitted by
Saransh Shivhare (200881),
Third Year Undergraduate,
IIT Kanpur (U.P)



CONTENT



Objective

Quick go through of the tool

Advantage over other similar Tools

Libraries Used

Future Scope and Potential
Improvements

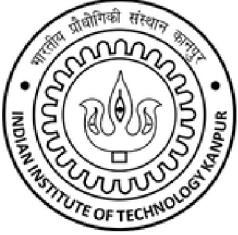
3

4

9

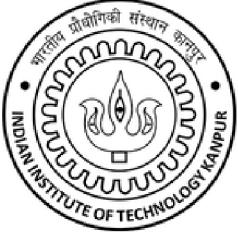
10

11



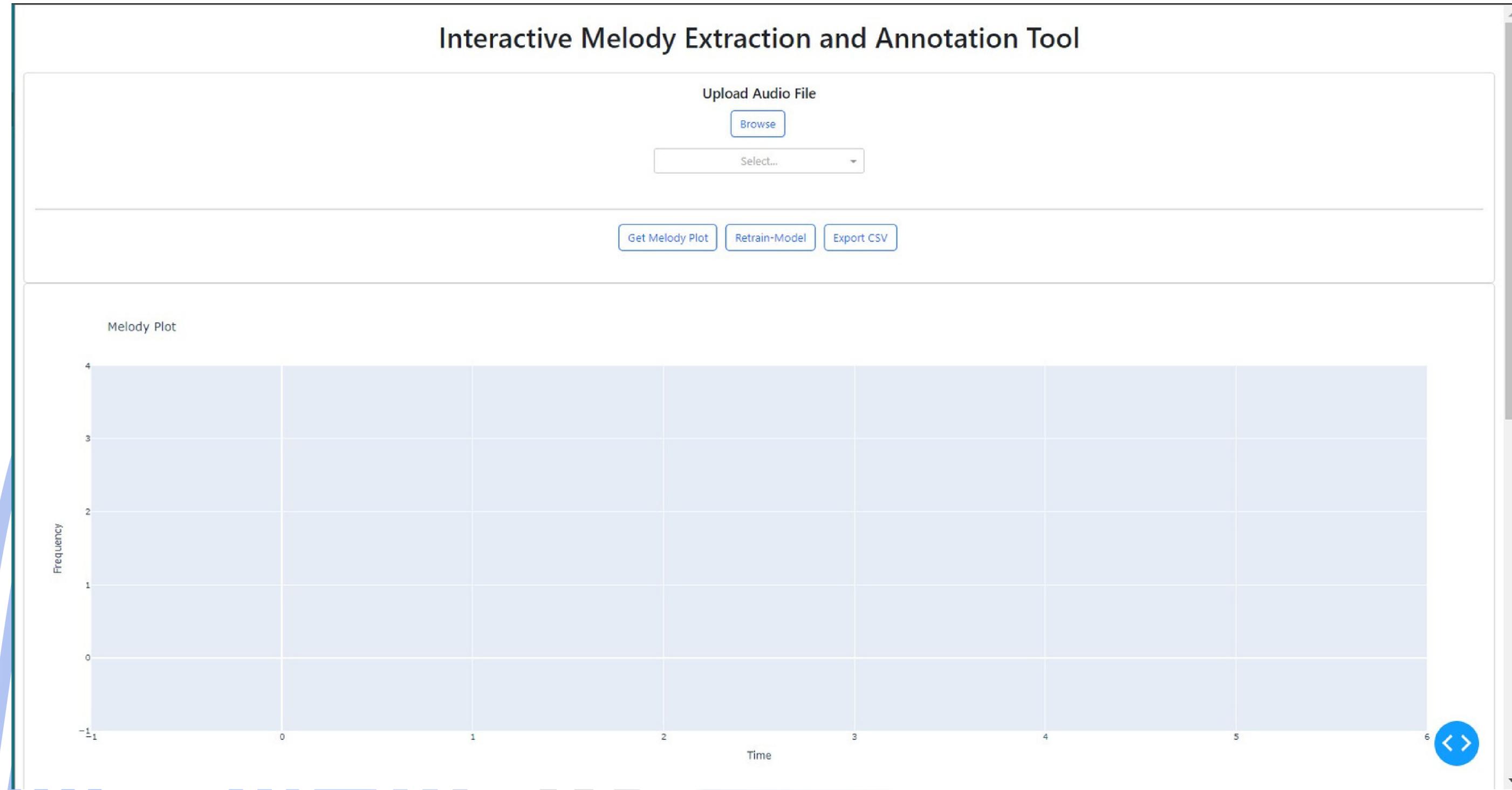
Overview

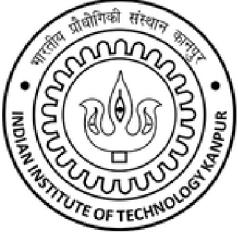
The goal of this project is to create an interface for the task of melody extraction and annotation. The tool is useful for creating a dataset for machine learning models related to music. This has applications in music transcription, music analysis, music recommendation system and many more related fields. Hence, building an interactive web platform to modify melodic data points predicted by the machine learning model is required. Also learned about some melody extraction and domain adaptation techniques.



EE392A (UGP-II)

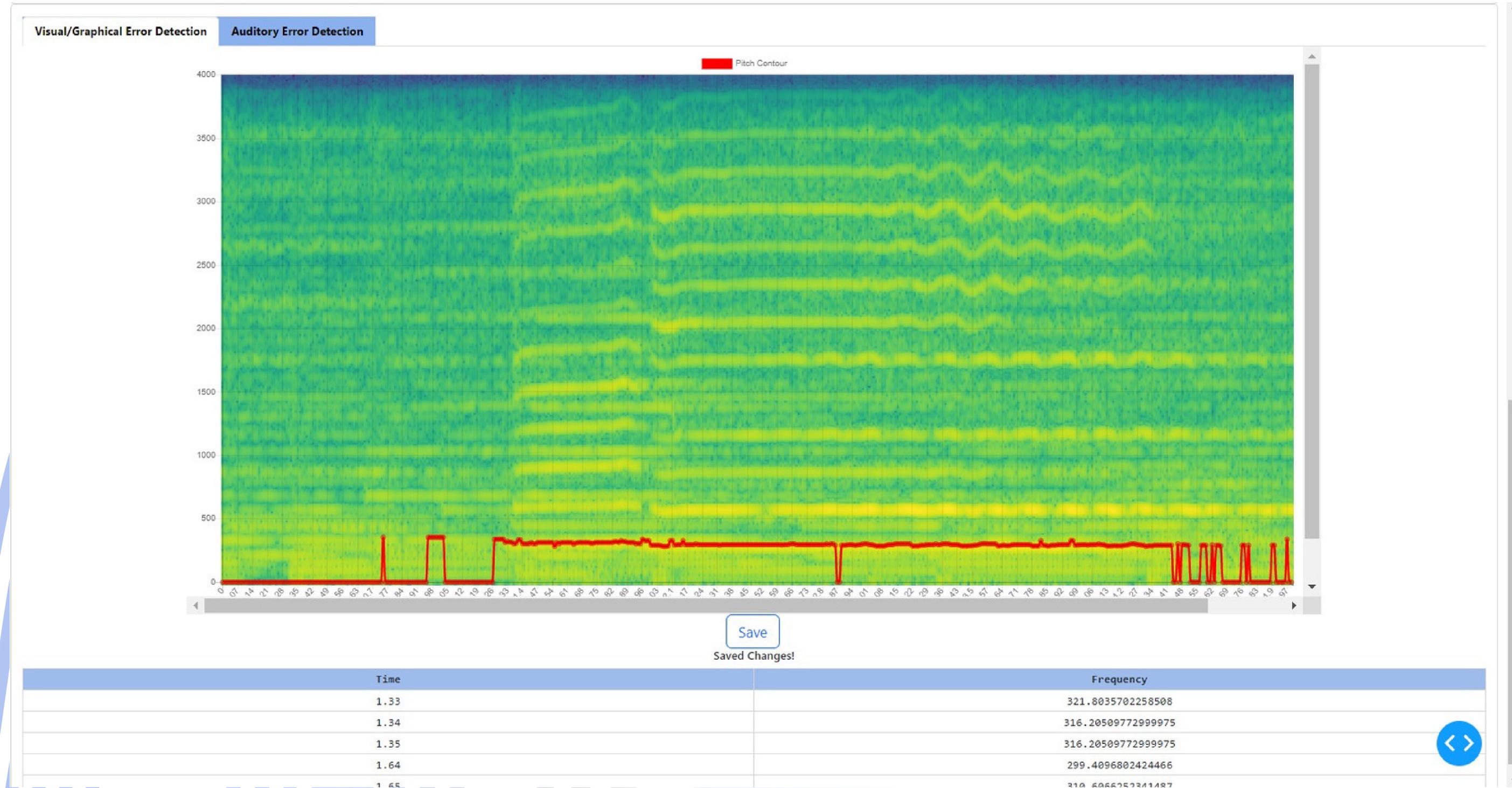
QUICK GO THROUGH OF THE TOOL

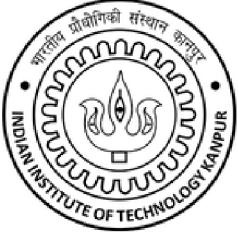




EE392A (UGP-II)

QUICK GO THROUGH OF THE TOOL





EE392A (UGP-II)

QUICK GO THROUGH OF THE TOOL

Visual/Graphical Error Detection Auditory Error Detection

Play Sine Wave Audio of desired frequency

Enter the frequency of Sine Wave you wish to Generate

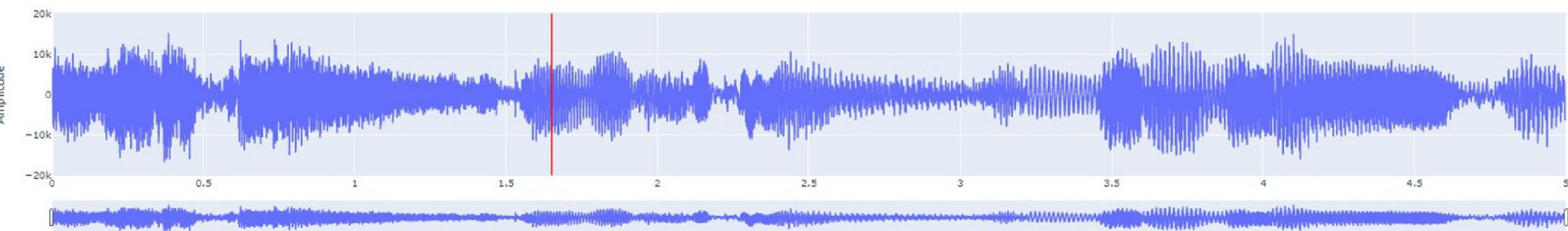
Generate Sine Wave Audio

▶ 0:00 / 0:03

Play Original Audio

Waveform Plot

Amplitude

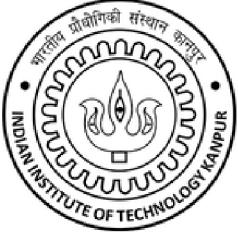


0 0.5 1 1.5 2 2.5 3 3.5 4 4.5 5

0:00 / 0:05

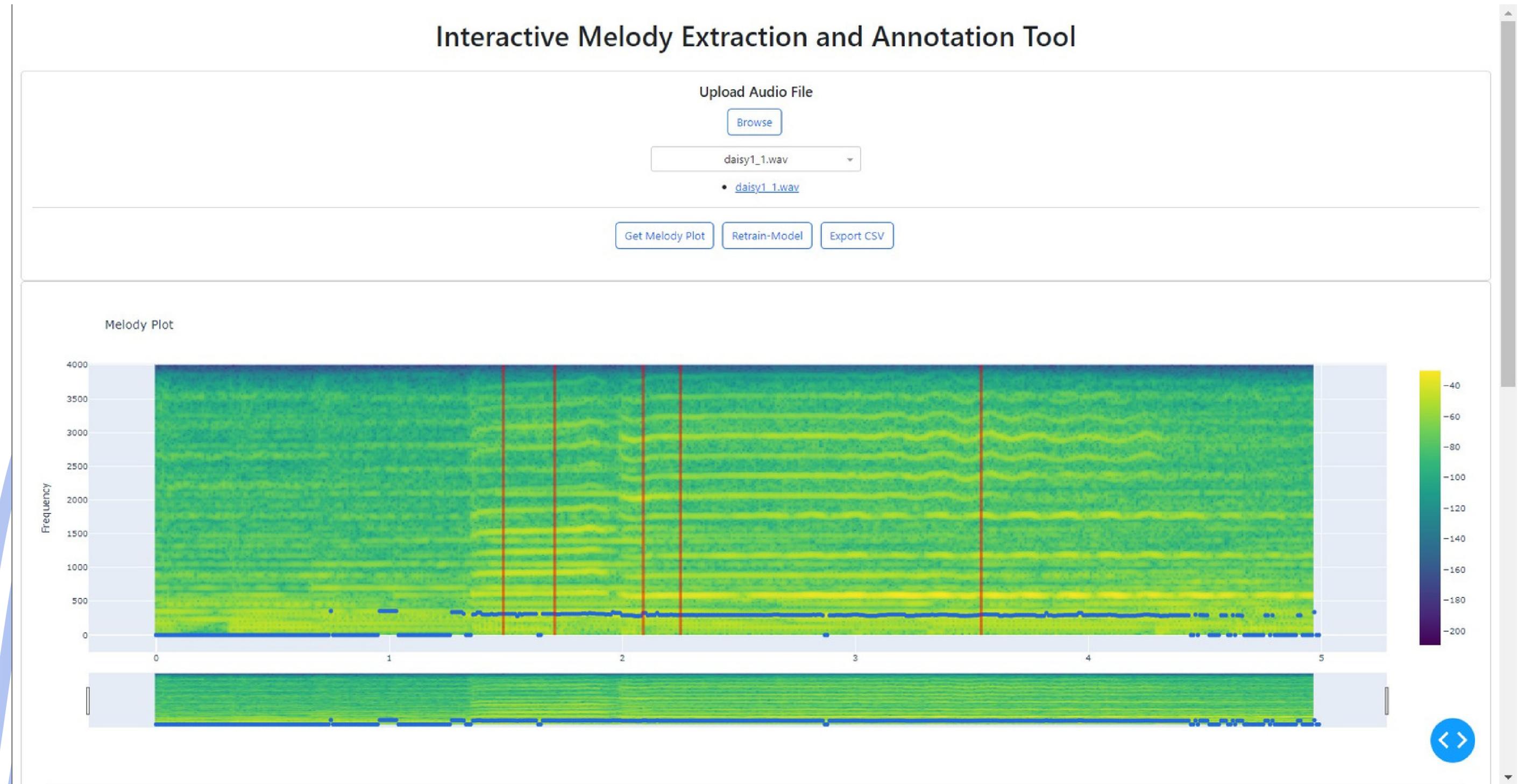
Get Wave Plot

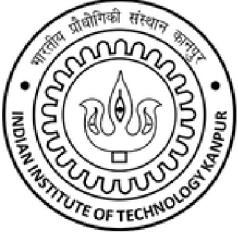
<>



EE392A (UGP-II)

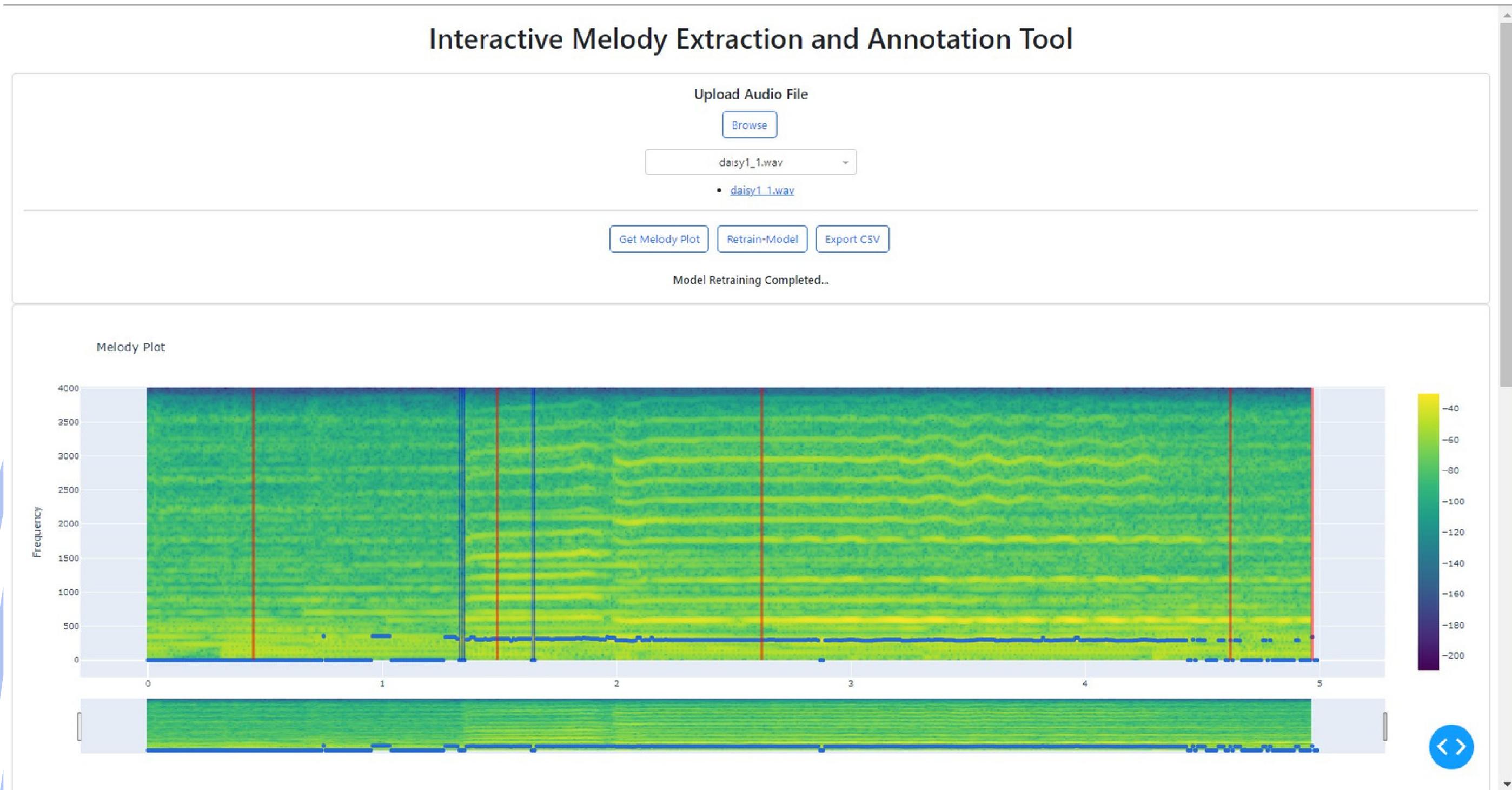
QUICK GO THROUGH OF THE TOOL

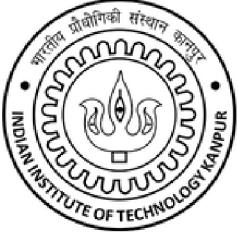




EE392A (UGP-II)

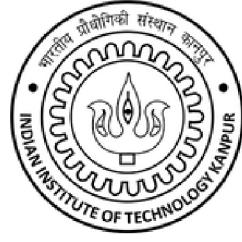
QUICK GO THROUGH OF THE TOOL





Advantage over previously available tools

- Users can annotate every predicted pitch point through the markers available in the interface.
- Users can detect errors in the predictions either by visual analysis or by auditory analysis.
- Low confidence regions are highlighted using red colour.
- Annotated data points are highlighted using blue colour.
- The machine learning model can be easily deployed in the backend. This will help the user to carry out visual examination of the model's performance.



Libraries Used

The programming language used in the application is majorly python and javascript. Some of the major libraries used in the application are:

DASH

Python library for developing
interactive dashboards

LIBROSA

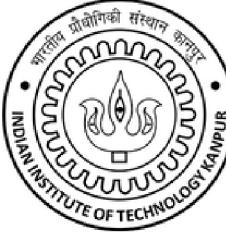
Python library for audio
processing related tasks

CHART.JS

JS library for interactive
charts and plots

PLOTLY

Python and JS library for
interactive plots

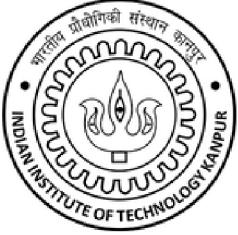


Future Scope & Potential Improvements

In the Auditory Error Detection section, [overlapping the melody audio file with the original audio](#) will make it easier for users to spot errors in the predictions provided by machine learning models. [This will allow for faster error identification.](#)

For a spectrogram of a particular audio, we aim to depict the least confident time frames by a separate color bar, so that visually it is easier to annotate those points. [This will assist the user in determining whether the performance of the model is increasing or decreasing over time.](#)

Currently, the web application is being executed on a client-side server, meaning it is [a client-side application](#). Hosting the application requires multiple adjustments to the backend's handling of the application.



EE392A (UGP-II)

End of Presentation...

Thank You