

Project Overview: Sound Refine

Brain and Cognitive Society

May 21, 2024

1 Mentors

- Srishti Chandra(221088)
- Deham Rajvanshi(220335)
- Devansh Gupta(220345)
- Aditya V(220082)

2 Motive:

Our final model is designed to process audio input contaminated with noise, removing the noise while transcribing the audio and potentially identifying the speaker. This functionality proves invaluable for tasks such as transcribing lectures or documenting unproductive meetings efficiently.

3 Problem Statement:

Our project involves Noise Cancellation and Audio Transcription. We aim to create a model that processes a mixed input of **clean speech** mixed with **background noise**, employing **machine learning algorithms** to effectively isolate and remove the noise, thereby reconstructing the **clean speech signal**. Subsequently, it transcribes this **purified speech** into a **textual form** in **Hinglish** language, to facilitate further analysis, understanding, or application in various domains such as communication, accessibility, and data-driven decision-making.

4 Project Details

The project involves exploration of practical applications of Deep Learning. The project focusses on mentees learning about Convolutional Neural Networks(CNNs), Artificial Neural Networks (ANNs), and Transformer Architectures. Additionally, mentees will delve into the theory and practical application of auditory processing techniques, including the exploration of relevant libraries. Key learning objectives include understanding CNN and Transformer Architectures, as well as mastering audio processing fundamentals. The mentees will be given resources and tasks related to it at the end of every week which they need to complete for the smooth conduction of the project.

5 PreRequisite

- Familiarity with some basic machine learning concepts.
- Proficiency in Python.
- Basic idea of libraries like NumPy and Pandas.
- Enthu and never-ending willingness to learn about ML.

6 Project Timeline

Week	Activities
Week 1	<p>Review of Basics</p> <ul style="list-style-type: none"> • Focus on basic Machine Learning techniques and introduction to Neural Network and Deep Learning. • Exploration of important libraries and frameworks such as TensorFlow. • We will be taking a quiz on the basics of ML and frameworks we have taught them in the week.
Week 2	<p>Introduction to Audio Processing Fundamentals</p> <ul style="list-style-type: none"> • This week aims to teach them about the theory of audio processing and introduction to libraries such as Librosa which is essential further in the project. • We will give them a task of data preparation based on Audio basics which they have learned. We will be using this dataset further in the project. • The data processing part involves mixing of clean speech and noise (cat noise, dog noise) for preparing the unwanted speech upon which we will be implementing our model.
Week 3	<p>Audio Processing libraries and Intro to CNN</p> <ul style="list-style-type: none"> • We would be continuing further on the Audio Processing Part and also give them introduction to CNN which we would be using further in our project.
Week 4	<p>A mini project of Instrument Detection.</p> <ul style="list-style-type: none"> • This week would involve building a mini project of Instrument Detection. This involves identification of which instrument is producing which sound. • This will help mentees relate CNN and audio which would help them further in the project.
Week 5	<p>Building our denoising model</p> <ul style="list-style-type: none"> • This week would involve teaching them about U-net architecture which we would be using our approach for denoising. • Using this mentees would learn about image deblurring which further would be extended based on the audio processing knowledge they have learned and the data set prepared to make a model of noise cancellation.
Week 6	<p>Learning about transformer architectures.</p> <ul style="list-style-type: none"> • This week would involve giving mentees an introduction to transformer architectures and hugging face library which would be helpful for Audio Transcription part of our project. • We would be giving mentees a text summarization task at the end of week for their understanding of transformers.
Week 7	<p>Building the Audio Transcription model.</p> <ul style="list-style-type: none"> • We would be teaching mentees about the 'Whisper' architecture which would involve building the Audio transcription model. • We would output text in 'Hinglish' language which is more prominently used.
Week 8	<p>Finetuning and Testing</p> <ul style="list-style-type: none"> • Comprehensive fine-tuning of the model. • Rigorous testing to ensure effectiveness and reliability.

7 Deliverables

1. **Quiz:** The mid-term evaluation will include a comprehensive quiz assessing the understanding and retention of foundational concepts taught during the first week.
2. **Data Preparation and Processing:** The second-week assignment involves cleansing, formatting, and organizing data to ensure its suitability for further analysis and modeling in the project.
3. **Instrument Detection Model:** Development of a model capable of accurately identifying instruments within a given dataset will be done by the mentees during the third week.
4. **Mid-Term Presentation:** Mentees will showcase their progress through quizzes and models they have created till now.
5. **Final Model of Noise Cancellation:** Mentees will finalize the noise cancellation model, aiming for optimal performance and effectiveness.
6. **Conversion into Hinglish:** The weeks evaluation includes the successful conversion of clean speech into Hinglish.
7. **End-Term Evaluation:** Mentees will present their final model that cancels noise and transcribes the audio present in Hinglish to text. The model will also have a proper user interface created by Gradio. If time permits then, we will also add the feature to recognize the speaker in this model.