MelodyMate

Harry's House

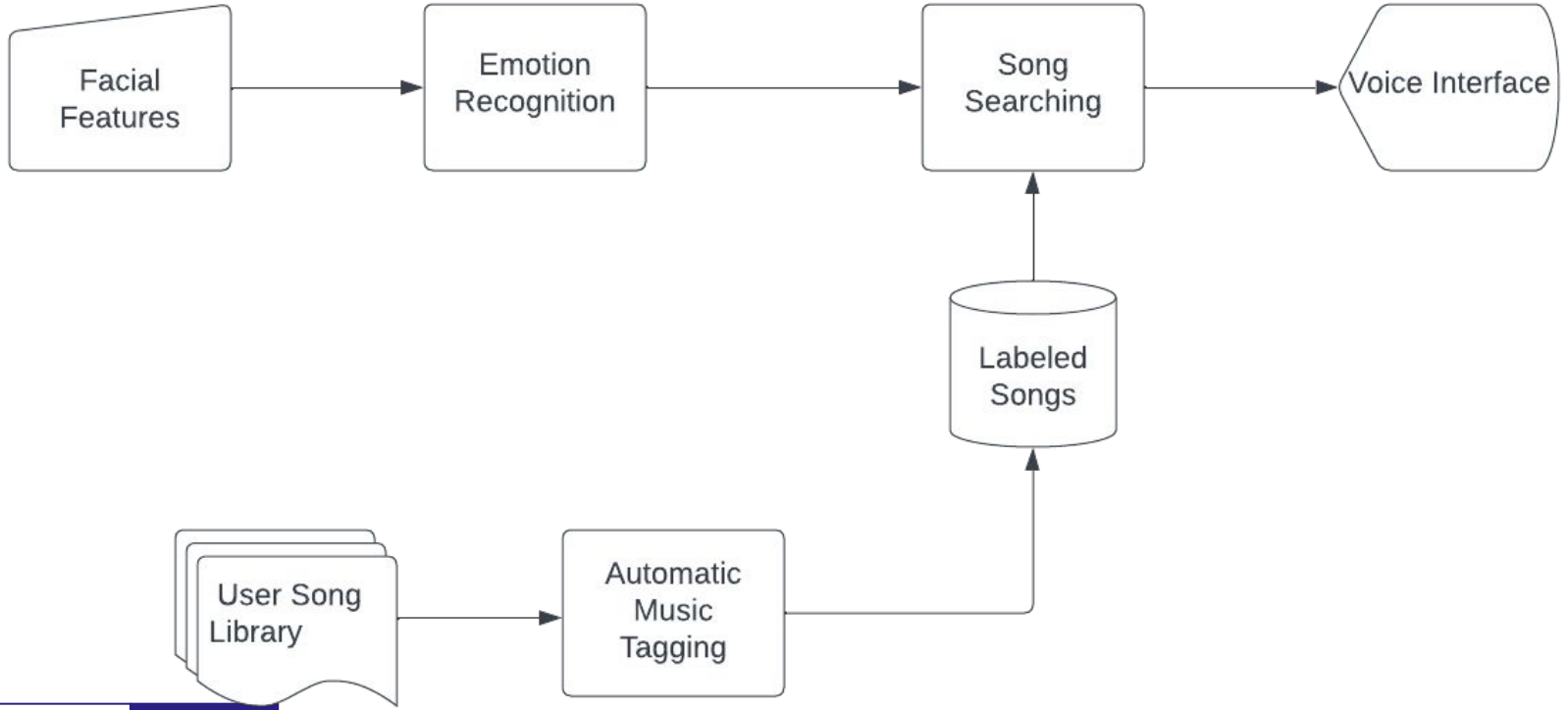
Harry Chen, Aditya Gutti, Rohan Gaikwad, Richard Wu
Bhargav Panguluru (TA)

Project Objective

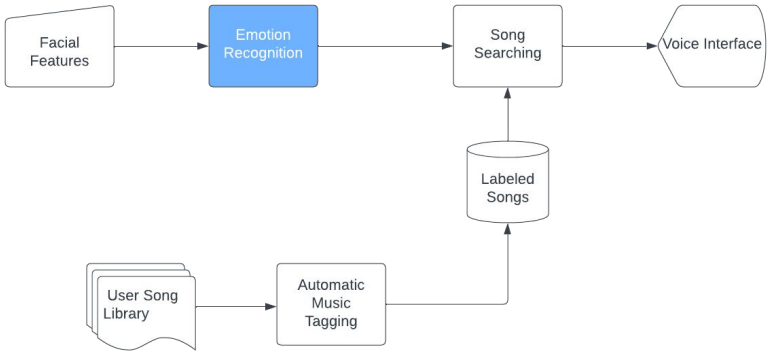
A virtual assistant that recommends songs based on the user's emotional state.

Sat Aug 5 9:48 PM

System Overview

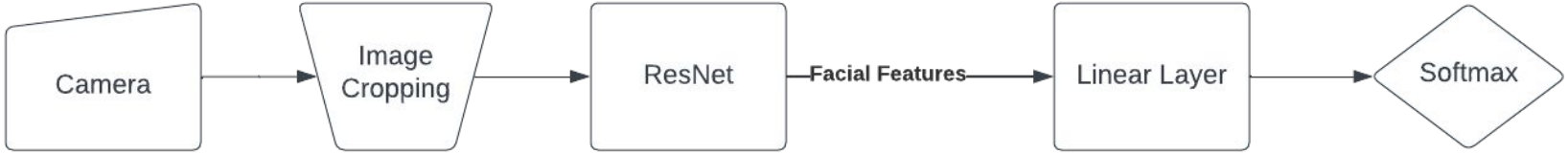


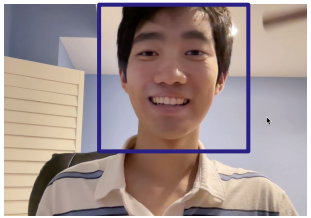
Emotion Recognition

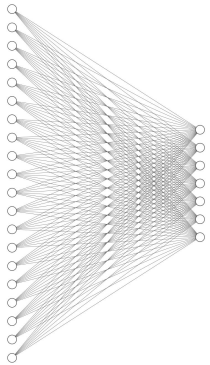


Model Overview

- Goal: classify a face into 1 of the 7 emotional states







512 -> 7

anger	0.02
happy	0.8
sad	0.05
...	...

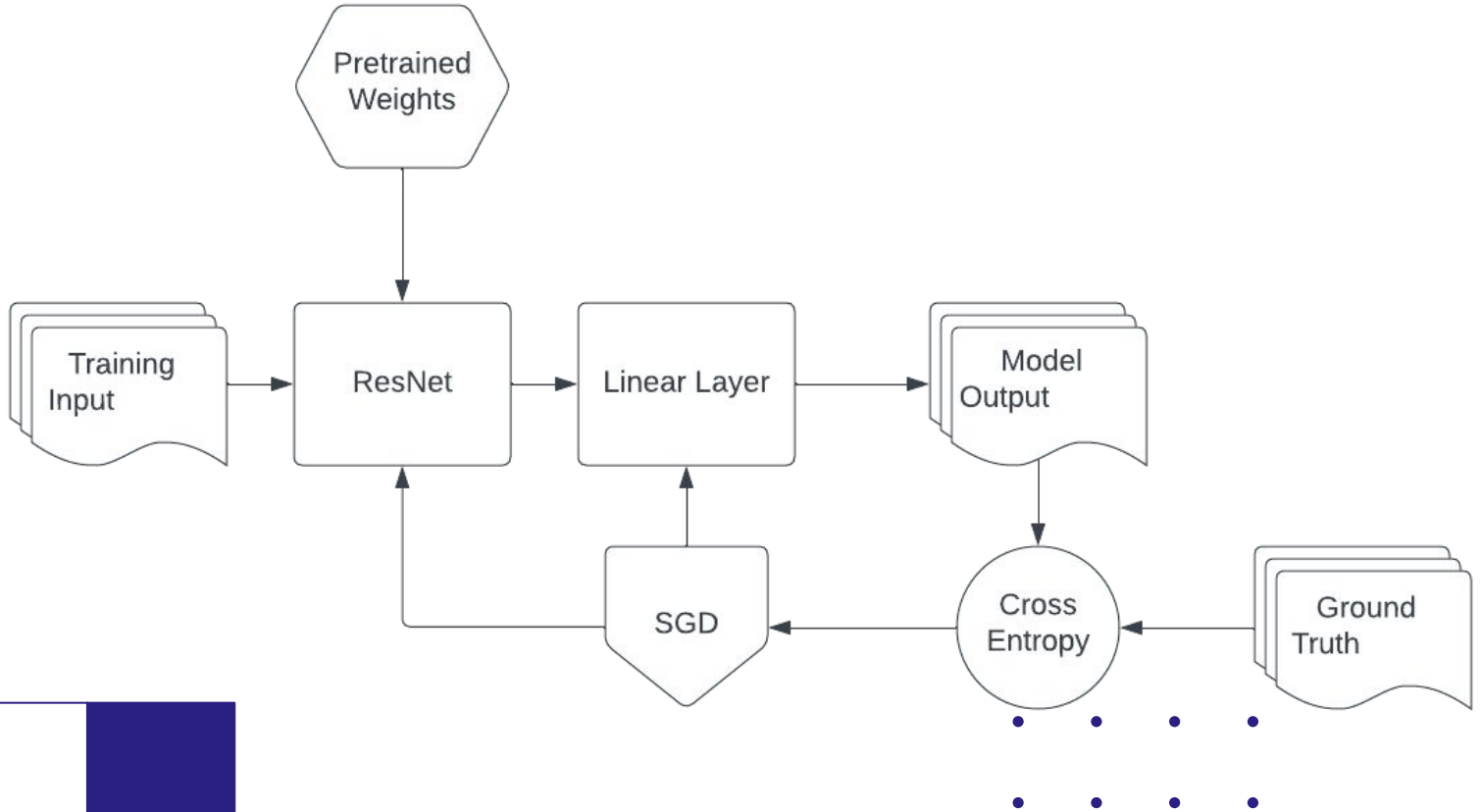
Data Set

- Emotion and identity detection from face images

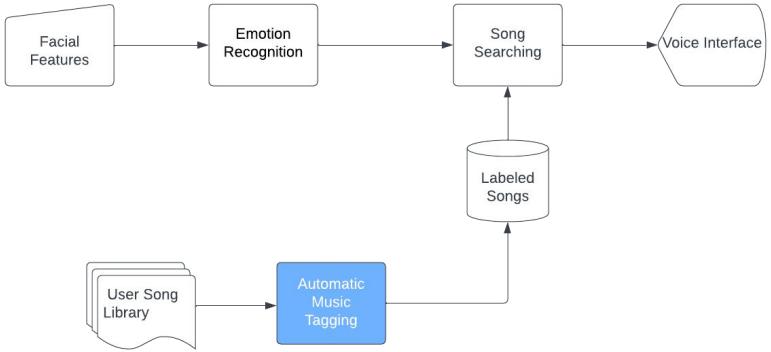
  (https://www.kaggle.com/datasets/noamsegal/affectnet-training-data)

  - ~1300 224 x 224 grayscale images labeled with 1 of the 7 emotional states

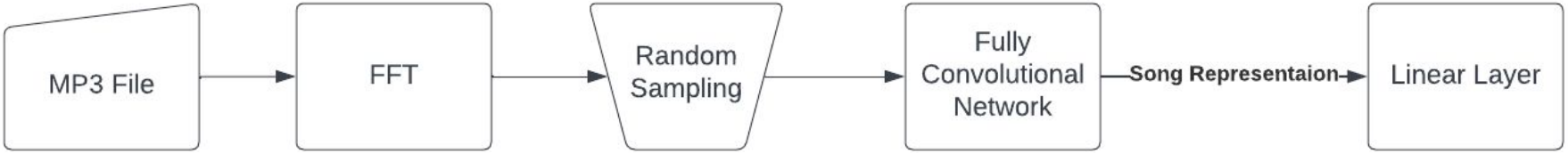
Training Process

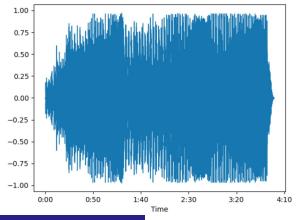


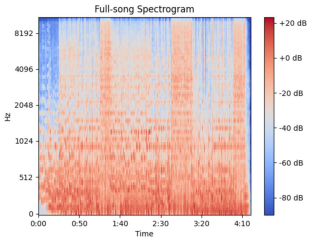
Music Tagging

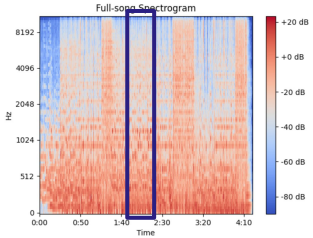


Model Architecture







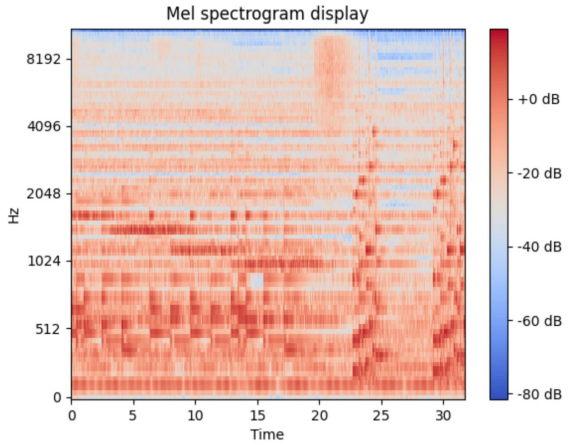


soft	0.5
calm	0.8
dark	0.05
...	...

No softmax due to multilabel

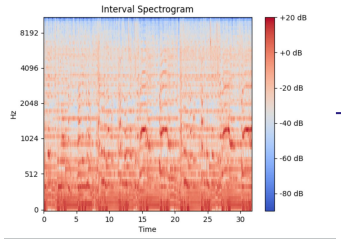
Data Set

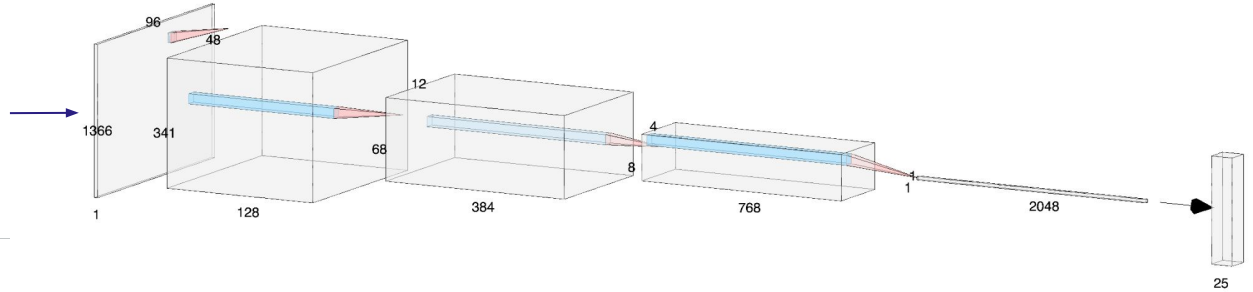
- MTG-Jamendo
  - ~12,000 tracks with labels for 25 tags/moods



[soft, calm]

Model Architecture





Spectrogram

- 30s

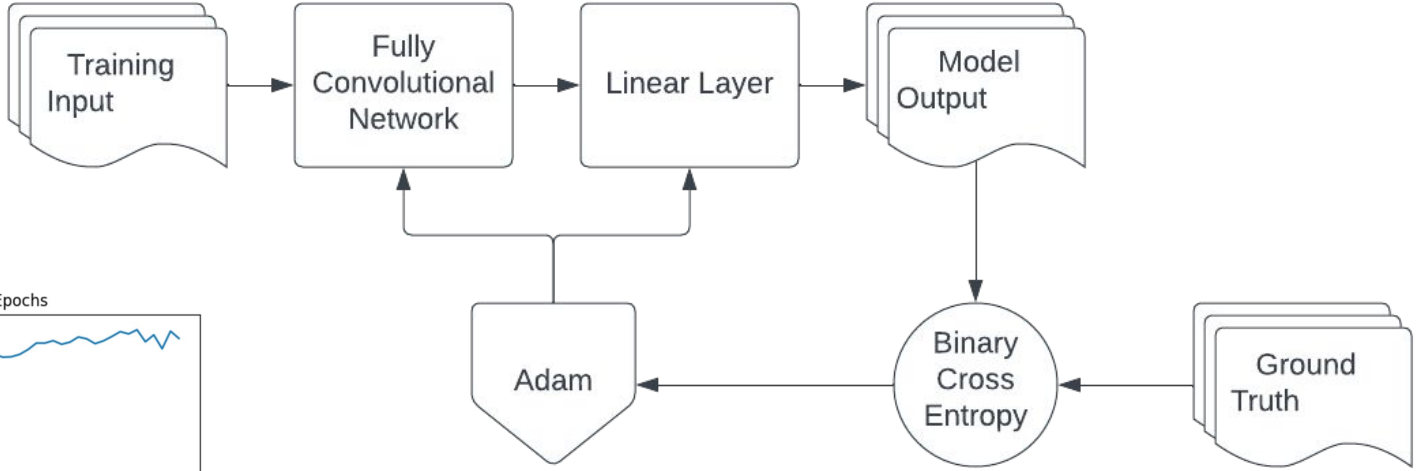
Fully Convolutional Network

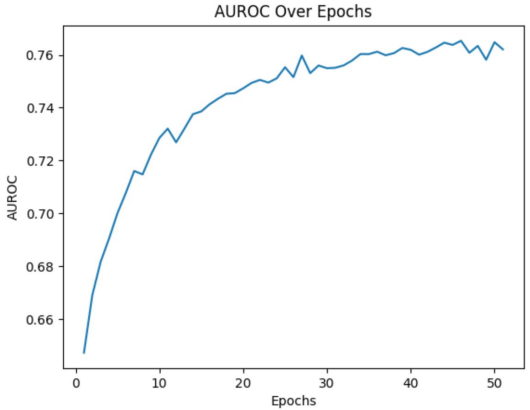
- ReLU

Linear Layer

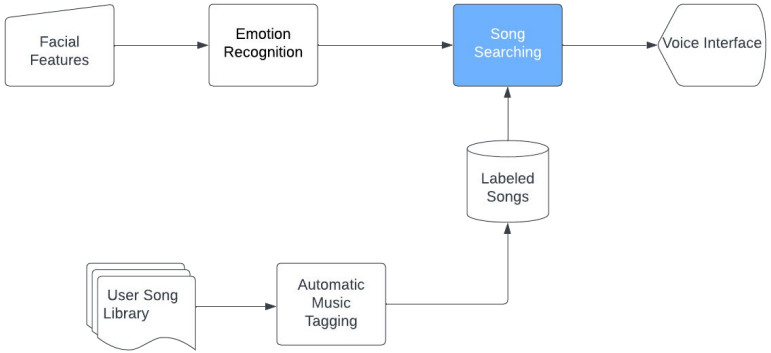
- Sigmoid
- Dropout

Training Process

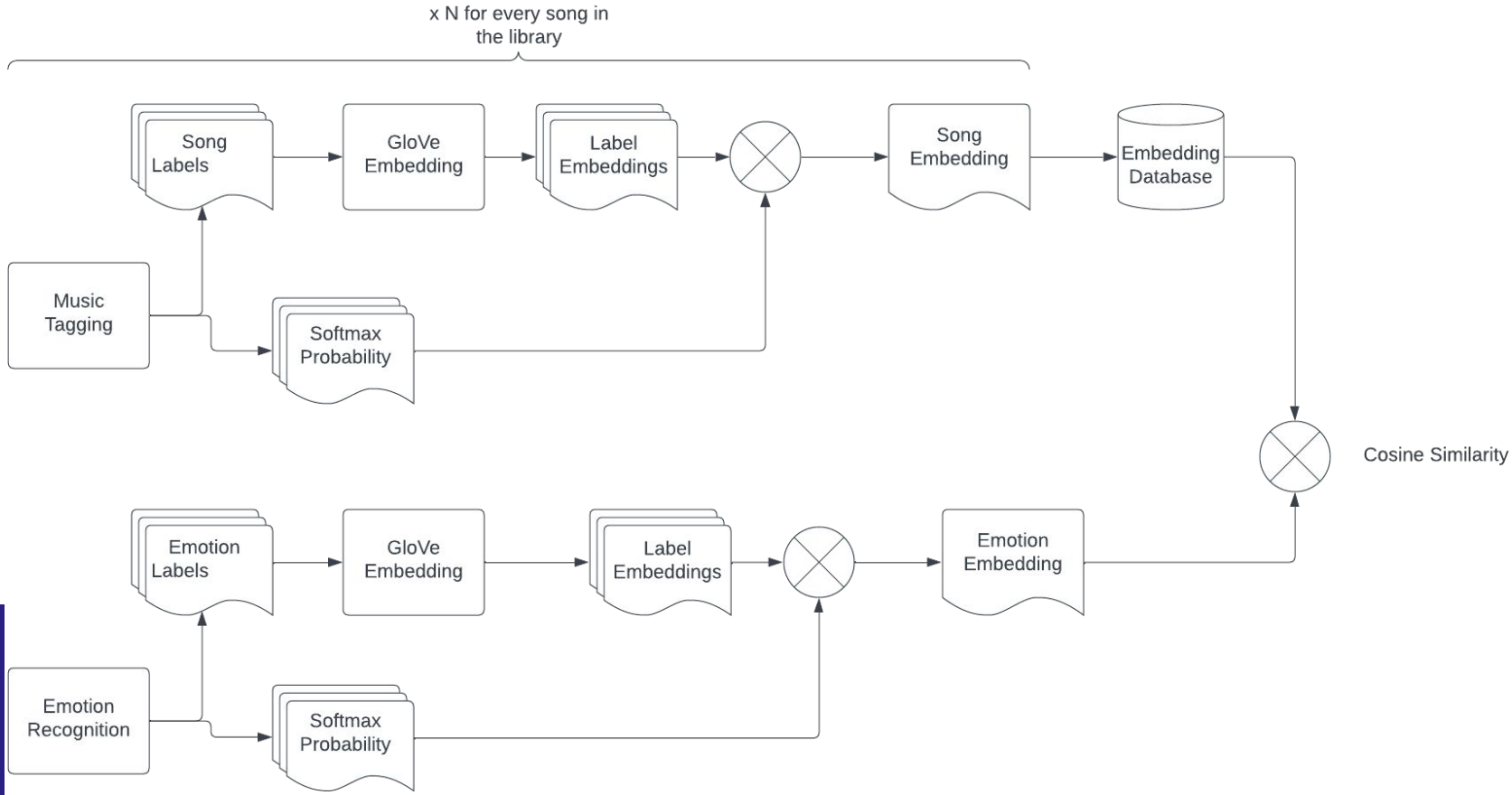




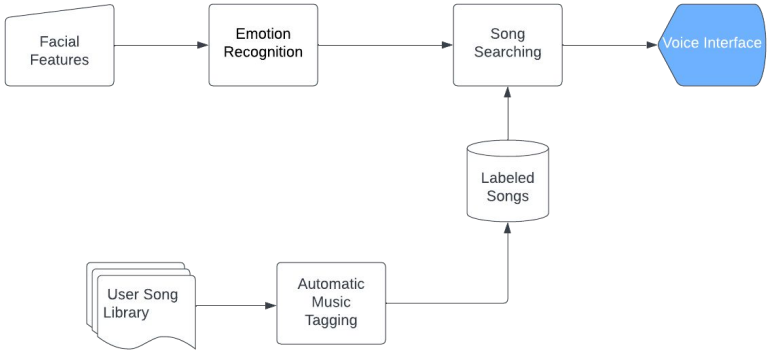
Song Searching



Model Architecture



Voice Interface



Voice Interface

- Speech-to-Text
- Text-to-Speech
- Prompt-engineered ChatGPT

ChatGPT

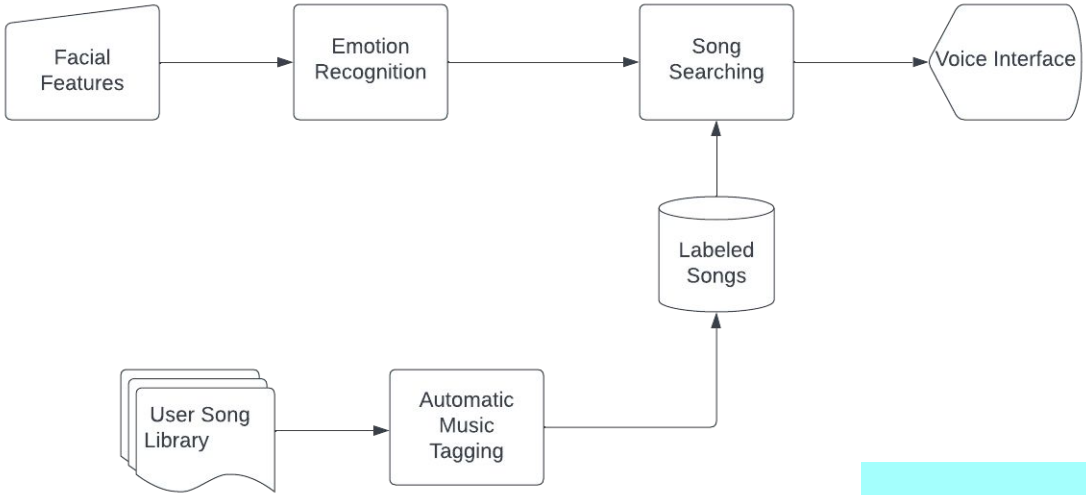
Future Goals

In the Future

- Attention-based music tagging for better accuracy
- Vision-transformer for emotion recognition
- Genre Clustering to filter for more specific user inputs
- Preload more songs into the personalized library

MelodyMate

A cognitive assistant that can recommend songs based on emotion



Questions?