Songs Generator App

Under the Supervision of :

Dr. Nada Ghneim

Eng. Massa Baali

A Senior Project by:

Amjad Almrai | Aous Abdulhamid

Table of Contents

01 Introduction

Music has always been tied to technology.

Song = Music + Lyrics +Speech

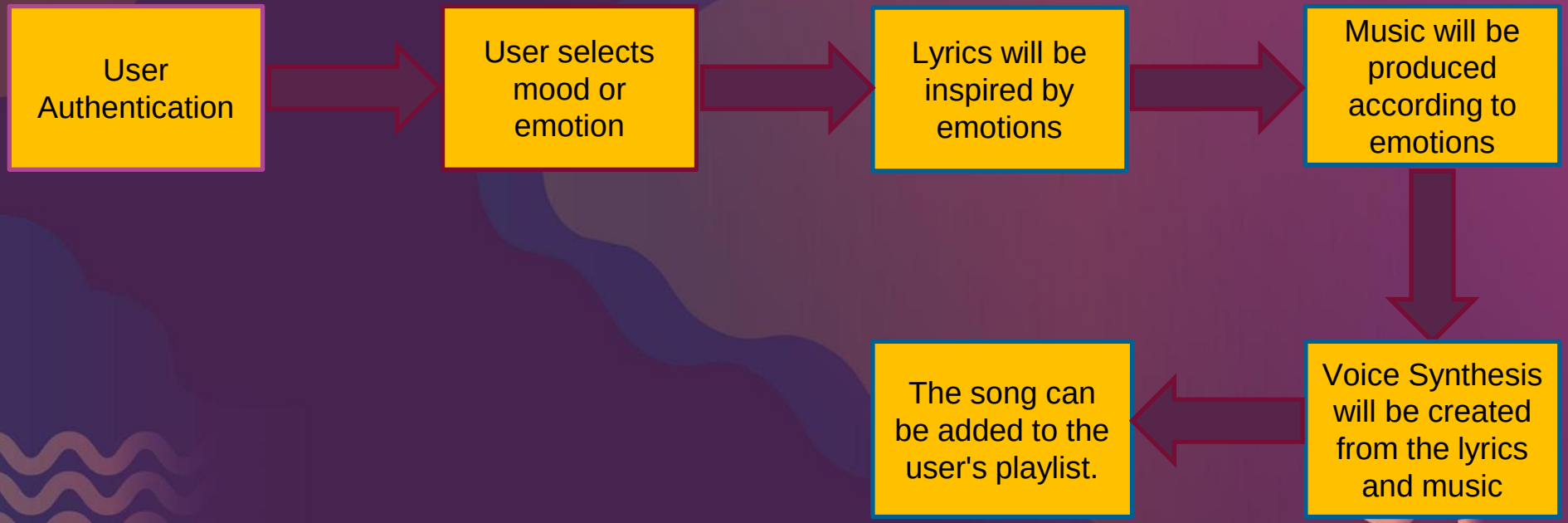
The creation of songs, music, and speech is a significant aspect of human creativity.
Can we teach machines to convey their emotions and ideas through AI songs

Soul Composer : is an app that generate lyrics and music according to the user's selected mood (happy, sad, etc.),then generate voice based on the lyrics and music that our app developed.

Comparison

	Boomy	My Lyrics Maker	Melobytes	Sound raw	Soul Composer
Lyrics Generator	X	From Topic	According To Your Lyrics	X	✓
Music Generator	✓	✓	✓	✓	✓
Voice Synthesis	Record Your Voice	X	✓	X	✓
Mood Choice	✓	X	X	✓	✓
Save As Play List	X	✓	X	✓	✓

Project Pipeline



Lyrics Generator

Structure

Model Objective: generate lyrics based on emotion

Dataset → Preprocessing → Architecture → Results

Data Set

- We have collected all of the songs from allmusic.com.
- Based on five emotions, five datasets
- has 200 songs, one for each emotion.

Preprocessing Data

01 Tokenization

02 Lowercase Words

03 Remove Punctuation

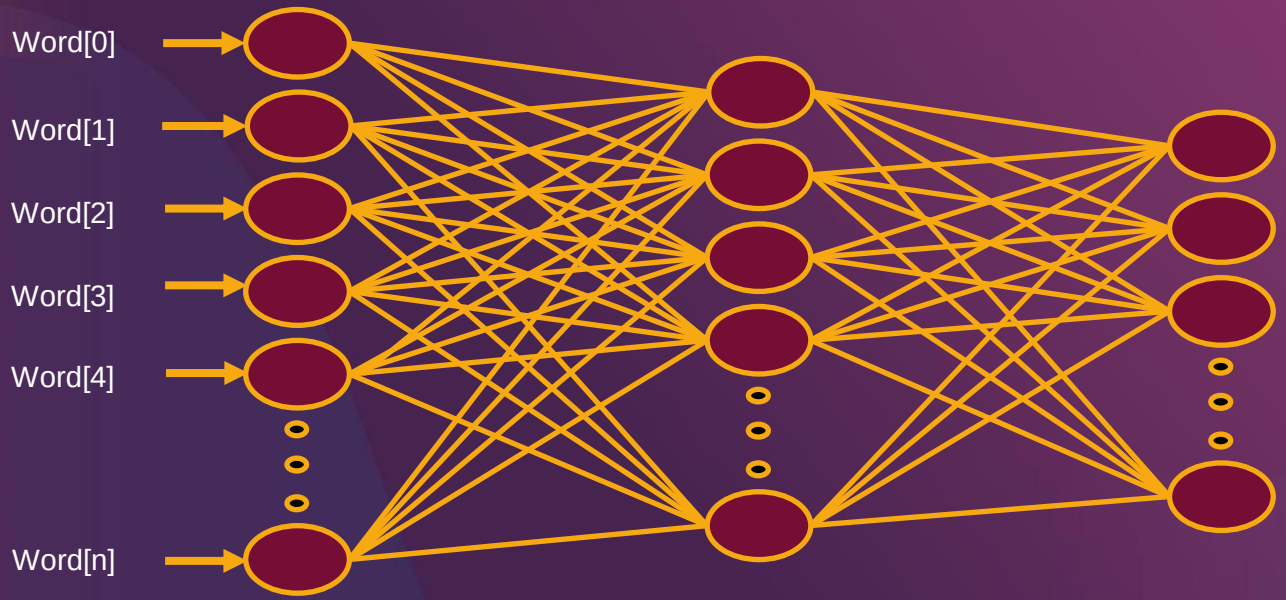
04 Remove Duplicates

05 Sequences Of Token

06 Sequence Of Integers

07 Separate The Sequence to input and output

Architecture



Input Layer **Hidden Layer** **Output Layer**

One Embedding Layer + two LSTM layers + two Dense layers with Relu

SoftMax Classification

HOW WILL THE MODEL DO IT?

4. Upon adding hyper parameters, lyrics is obtained.

3. The output of the LSTM was sent to two dense layers.

2. The output of the embedding layer is then taken by the two LSTM layers.

1. Word index and sequence length are entered into embedding layer.

Results

Hyperparameters :

- **Activation** : Relu and softmax
- **Different epochs** : for each five models
- **Batch size :** 128
- **optimization** algorithm : Adam
- **loss function** : categorical crossentropy

After applying hyper parameters to our five models, we obtain varying degrees of accuracy (92–96%).

Music Generator

Structure

Model Objective : generate music depending on emotion from midi files

Dataset → Preprocessing → Architecture → Results

Data Set

- Our dataset POP909 includes piano tunes stored in the MIDI (Musical Instrument Digital Interface) format
- Using MuseScore3, we separate that dataset to five emotions.
- Each dataset contains between 50 to 85 midi file
- music21 (a toolkit for computer-aided musicology, MIT) These MIDI files data were extracted using a Python toolbox.

Preprocessing Data

01 All dataset will be transformed into music21 objects.

02 Extract chords from objects scores

03 Obtain notes and their duration.

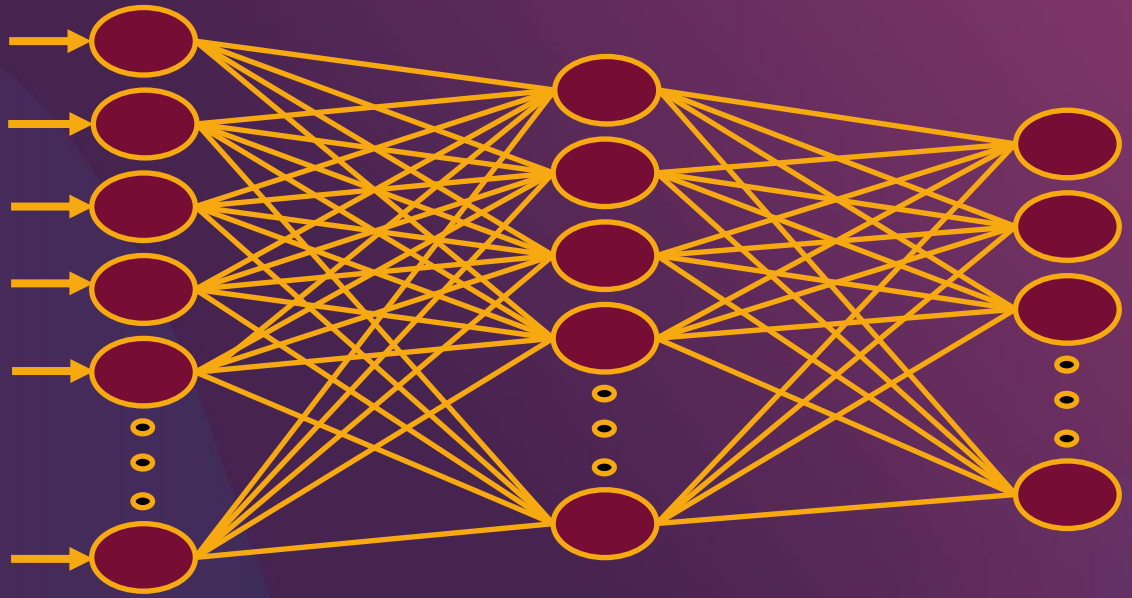
04 Choose the C major key

05 Mapping chords and durations to integers

06 Create Sequence Of Integers for chords and durations

07 *Construct train and target sequences for chords and durations*

Architecture



Input Layer

Hidden Layer

Output Layer

Two Embedding Layer + two LSTM layers + one Dense layers with tanh Function

SoftMax Classification

HOW WILL THE MODEL DO IT?

4. Upon adding hyper parameters, notes and durations are obtained.
3. Dense layer changing the output's dimension to 1
2. Then, the two LSTM layers take the concatenated versions of both inputs.
1. Notes and durations are entered into two embedding layers.

Results

Hyperparameters :

- **Activation** : Tanh and softmax
- **Different epochs** : for each five models
- **Batch size :** 128
- **optimization** algorithm : RMSprop
- Learning rate :0.01
- **loss function** : categorical crossentropy

After applying hyper parameters to our five models, we obtain varying degrees of accuracy (94–96%).

Voice Synthesis

Structure

Model's objective : is to generate voice from midi files and lyrics.

Hidden Markov Models (HMMs) were used to create a singing voice synthesizer.

How ?

Preprocessing the midi and the lyrics 01

02 Check The Midi File if Contains one Channel

Verify that the wording matches the amount of notes 03

04 Define the Voice's Speak Rate

Define the voice's tempo 05

06 Pass Mid, Lyrics, Speak Rate, and Tempo using Mid2Voice Models.

Call Voice Render in order to hear the voice 07

Data Set

Table 1: *Singing voice database.*

Singer	1 male (non-professional)
Songs	60 Japanese children's songs (about 72minutes in total)
Sampling Rate	44.1kHz
Quantization	16bit

Table 2: *Mel-cepstral analysis condition.*

Sampling Rate	16kHz
Frame Shift	5ms
Window Length	25ms
Window Function	Blackman Window
Spectral Feature	24 mel-cepstral analysis [8]

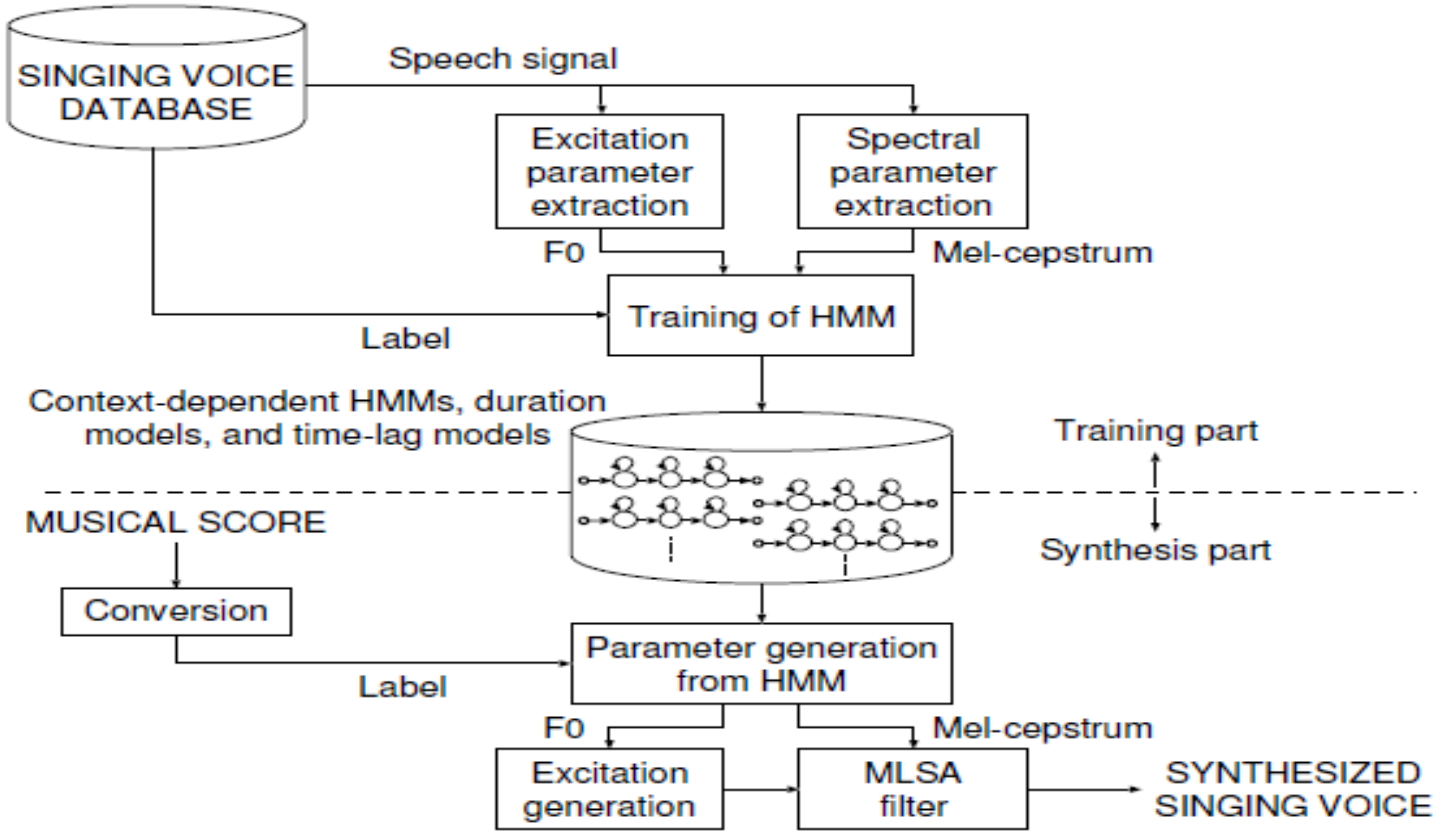
Voice Synthesizer

- In the training part, first we extract spectral and excitation parameters from a singing voice dataset and then they are modeled by context-dependent HMMs.
- Context-dependent state duration models and time-lag models are also estimated.
- In the synthesis part, first musical score to be synthesized is converted to a context-dependent label sequence.

Voice Synthesizer

- according to the label sequence, a song HMM is constructed by concatenating the context-dependent HMMs.
- spectral and excitation parameters are generated by HMM
- Finally, a speech waveform is synthesized directly from the generated spectral and excitation parameters using Mel Log Spectrum Approximation

Voice Synthesizer



Voice Synthesizer Example

We would want to express our gratitude to the doctors and listeners, in particular Dr. Nada and Dr. Masa, for guiding us during this journey.

05

Project Application

How We Do That ?

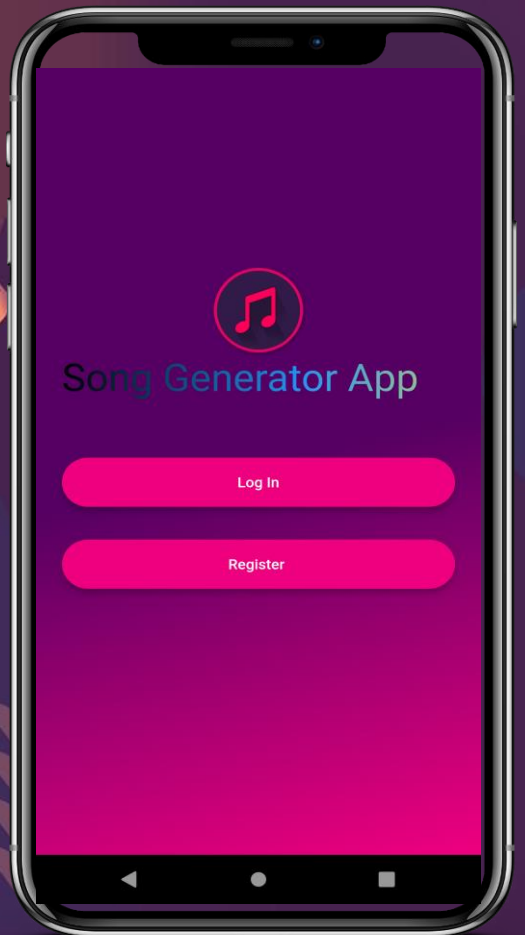
User selects lyrics and music emotion, and API is called.

When an API detects an emotion, it sends it to a model, which then creates music and lyrics.

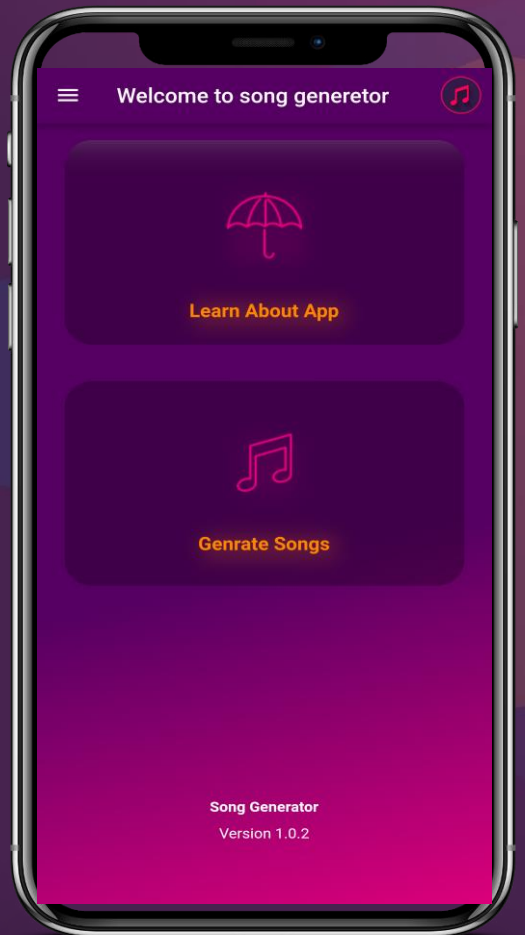
Generated music and lyrics are sent to the voice model. to generate a voice

Send music , lyrics and voice to the application, the user can now play it

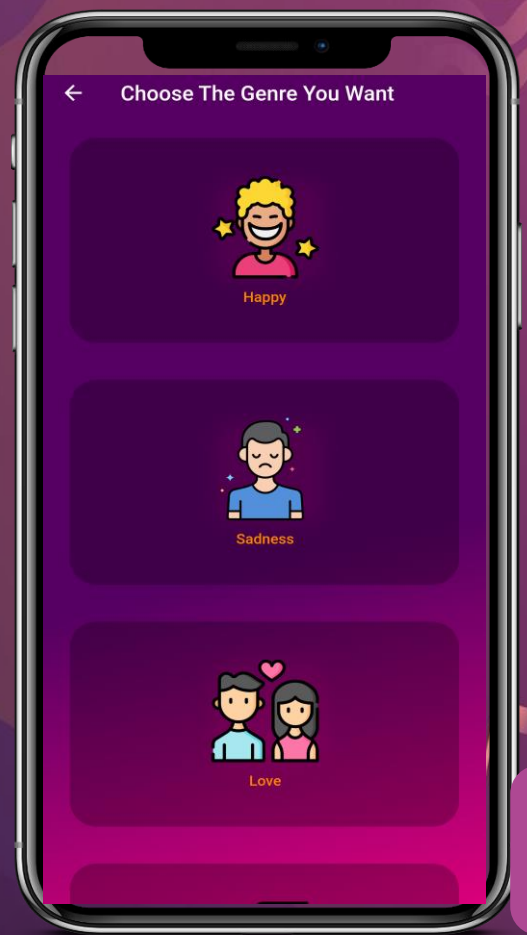
1


Song Generator App
Log In
Register

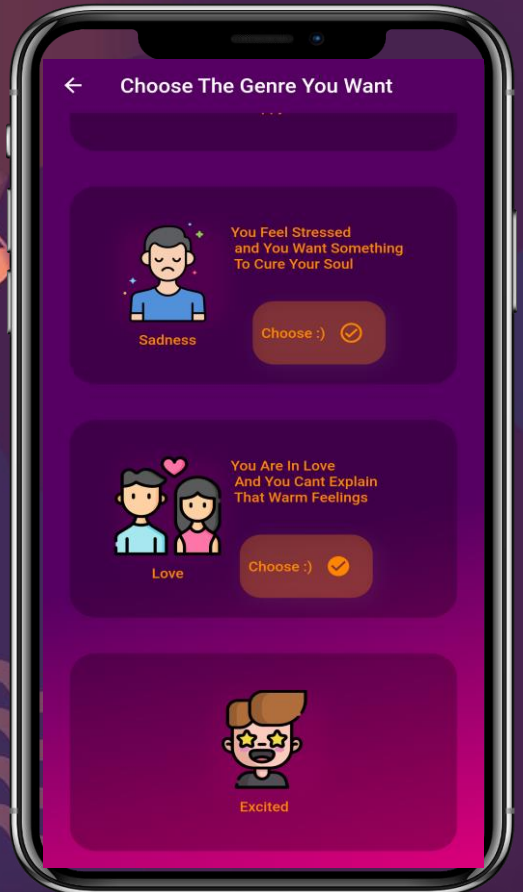
2


Welcome to song generetor
Learn About App
Genrate Songs
Song Generator
Version 1.0.2

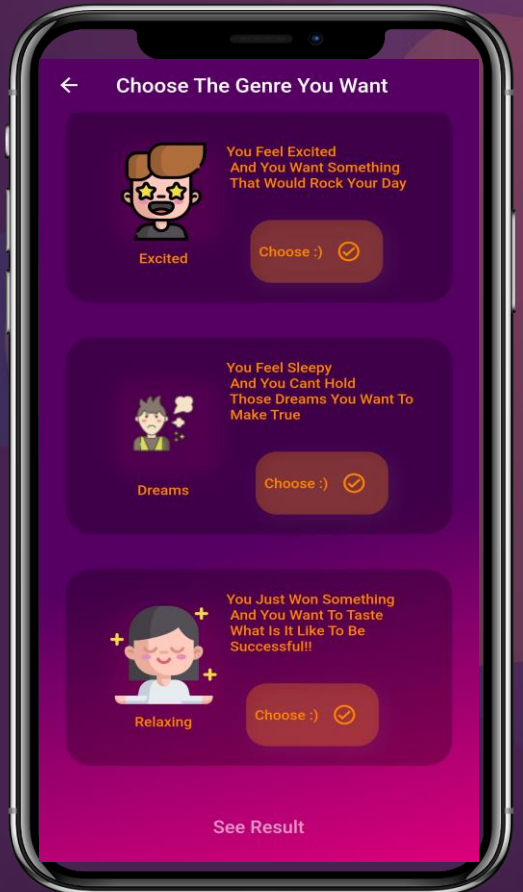
3


Choose The Genre You Want
Happy
Sadness
Love

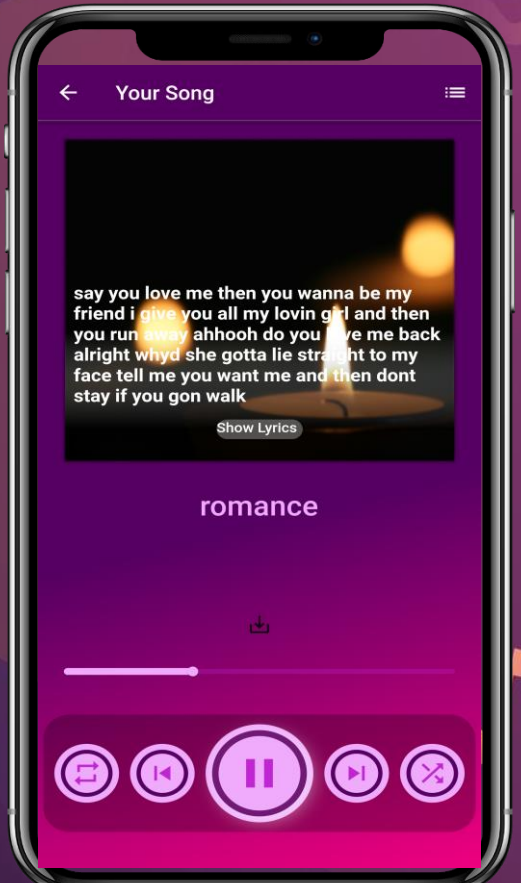
4


Choose The Genre You Want
You Feel Stressed
and You Want Something
To Cure Your Soul
Sadness
Choose :)
You Are In Love
And You Cant Explain
That Warm Feelings
Love
Choose :)
Excited

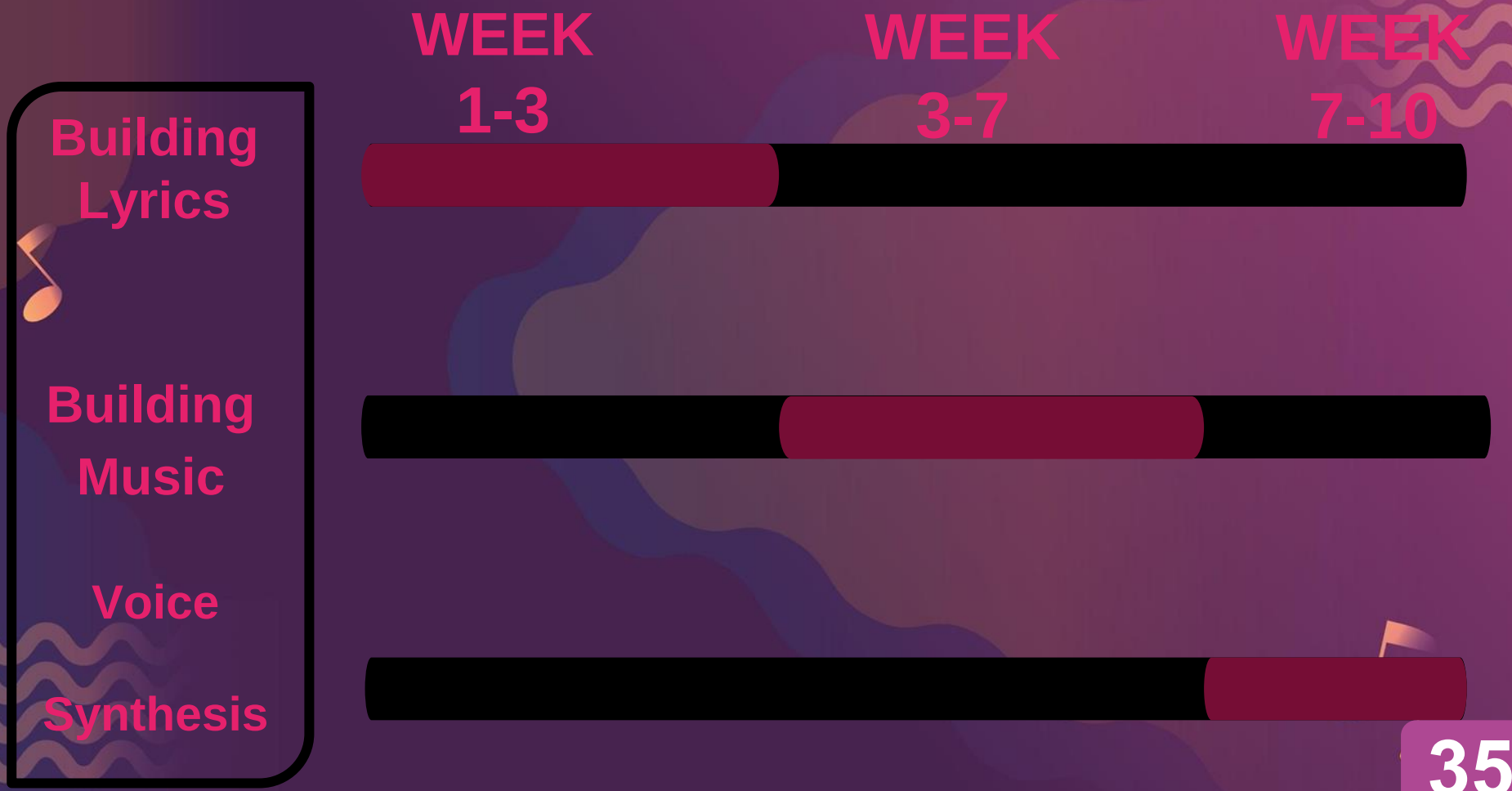
5

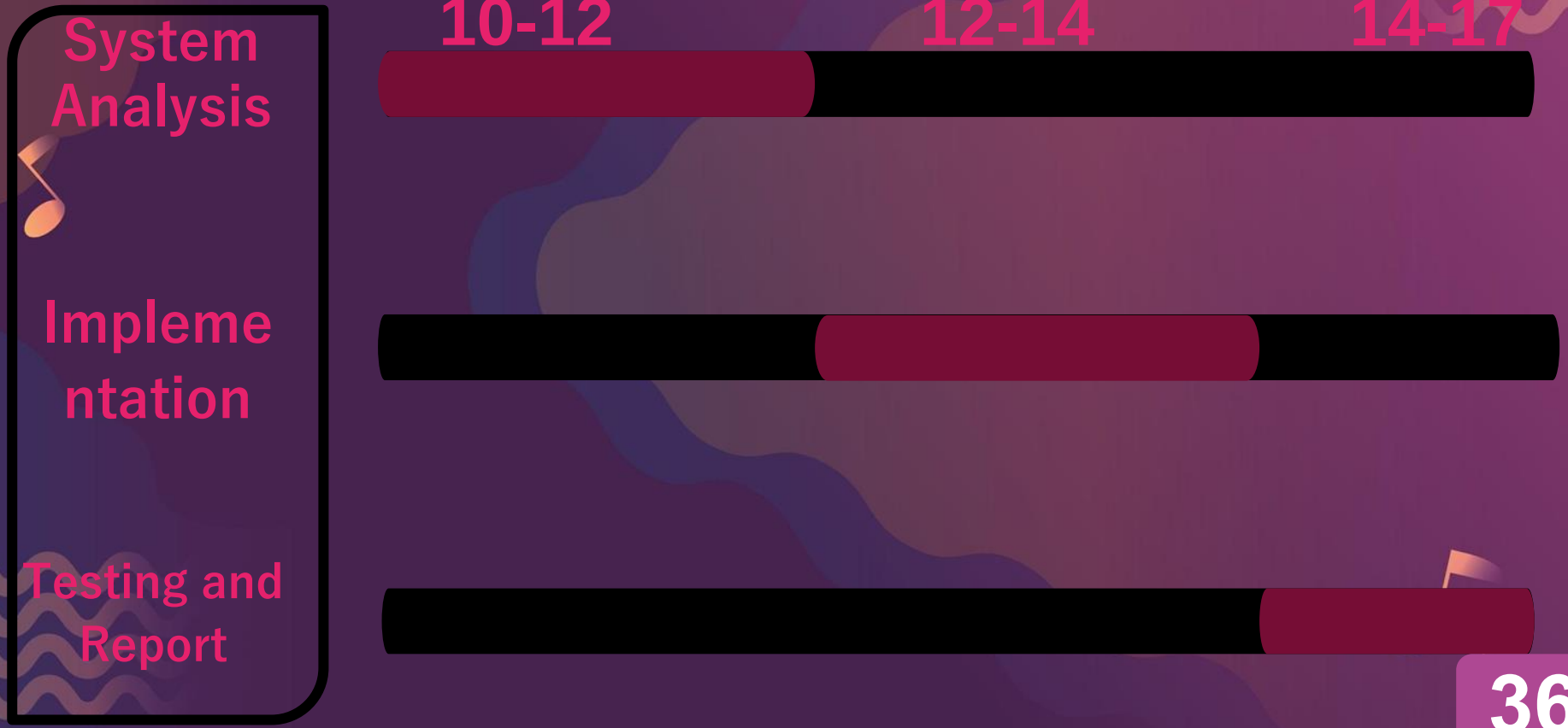

Choose The Genre You Want
You Feel Excited
And You Want Something
That Would Rock Your Day
Excited
Choose :)
You Feel Sleepy
And You Cant Hold
Those Dreams You Want To
Make True
Dreams
Choose :)
You Just Won Something
And You Want To Taste
What Is It Like To Be
Successful!!
Relaxing
Choose :)
See Result

6

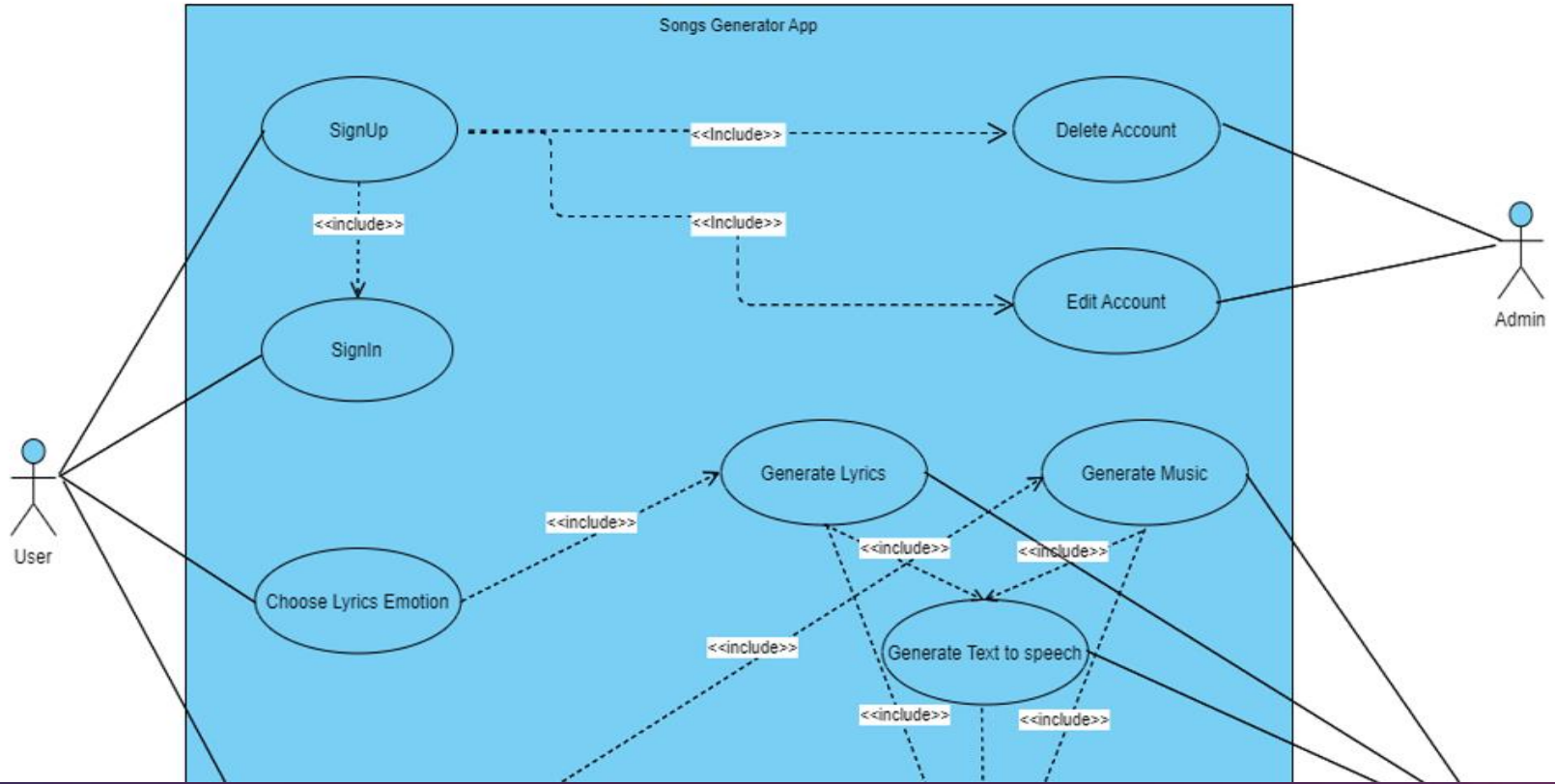

Your Song
say you love me then you wanna be my
friend i give you all my lovin girl and then
you run away ahhooh do you love me back
alright whyd she gotta lie straight to my
face tell me you want me and then dont
stay if you gon walk
Show Lyrics
romance

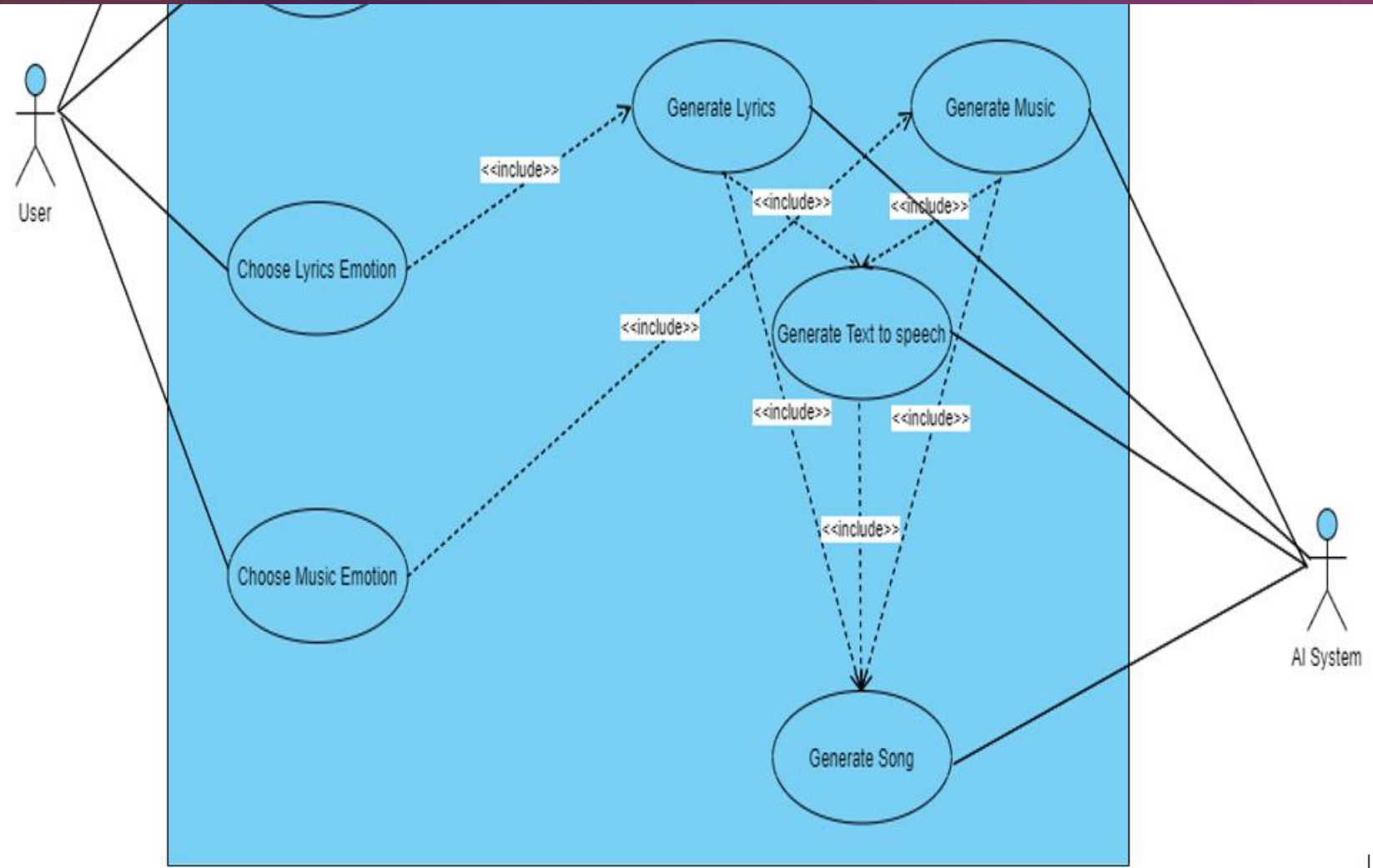
Project Timeline


WEEK 1-3
WEEK 3-7
WEEK 7-10
Building Lyrics
Building Music
Voice
Synthesis


WEEK 10-12
WEEK 12-14
WEEK 14-17
System Analysis
Impleme ntation
Testing and Report

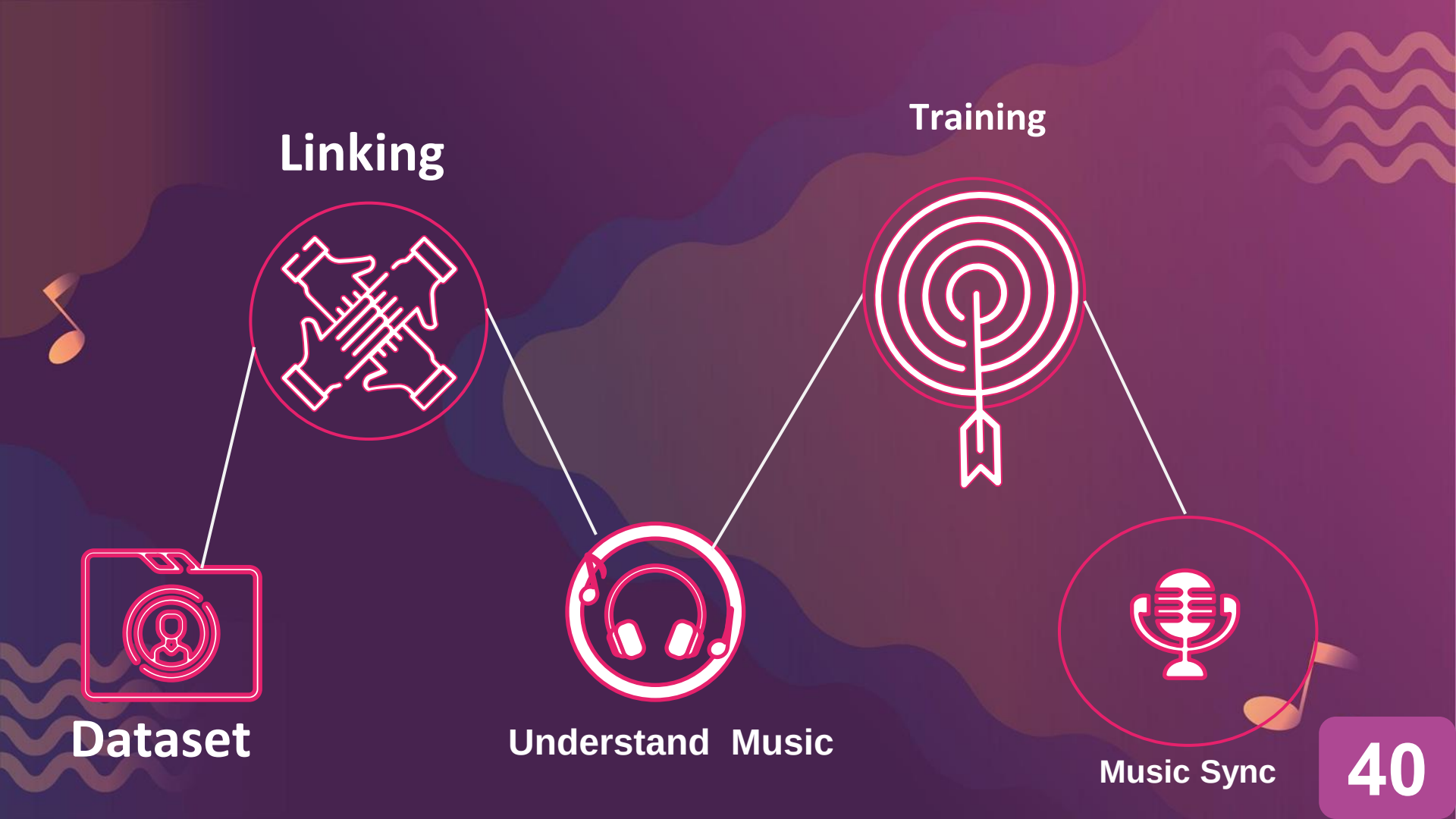
Use case Diagram




User
Generate Lyrics
Generate Music
<<include>>
<<include>>
<<include>>
Choose Lyrics Emotion
<<include>>
Generate Text to speech
<<include>>
<<include>>
Choose Music Emotion
<<include>>
AI System
Generate Song

07

DIFFICULTIES & FUTURE VISION


Linking
Training
Dataset
Understand Music
Music Sync

Song in another language

Voicing for particular artists

lyric video for the song

Various instrument music

08 Used Technologies

Flutter

POSTMAN

Flask

Application Building

Keras

Model Developing

THANKS!

For your
time, attention, everything!