COMP47700 PROJECT REPORT

CPR BEAT DETECTION FROM AUDIO

Thirumla Devi Sivakumar / 22200730

Note: The following are abbreviations introduced in this report:

- CPR - Cardiopulmonary resuscitation
- WAV file - Waveform Audio File Format
- sec - seconds
- ms - milliseconds
- mins - minutes

Introduction:

CPR is the basic life saving technique in case of emergencies where a person is unconscious. Usually the general public who are not health workers, trained or untrained, are advised to keep in mind a music/rhythm to which CPR is performed. A study by Alexander Muacevic and John R Adler on *"The Value of Songs for Teaching and Learning Cardiopulmonary Resuscitation (CPR) Competencies: A Systematic Review" (National library of Medicine, May 2021)*[1] supports this view.

Usually this music or beat suggestion would be in english. This project allows the general public to upload music or any audio and find portions of song/audio that can be used to perform CPR to, if it exists. This will allow people from all backgrounds and not only English speakers to know music from their native language that can be kept in mind while doing CPR for ease.

This application can be further extended to provide CPR music suggestions by language of users choice, so that they'll have an idea on what type of music/rhythm that may be used to do CPR.

Objective of the application:

The objective of this project is to come up with the optimal and fastest strategy to find the following and the challenges that constraint from doing so:

A. If the song is CPR performable
    a. If, yes then what all portions of the song can be used to perform CPR
    b. If, not what is the bpm of the audio that the song got rejected

B. Implement and analyse different strategies to check for CPR doable portions of an audio.

C. Compare and evaluate the correctness of the results from the above strategies through subjective and objective tests

D. Further assess any noticeable changes in results of subjective and objective tests. Also compare these results with the initially taken expert reference dataset.

E. Record if any interest biases found during subjective analysis.

Reference dataset used:

For the purpose of this report, the following audios were run through this project and used for subjective tests. The audios are from different languages like English, Spanish and Korean. The reference on whether this particular music is CPR-able or not was determined by 5 experts, 2 trained doctors, 1 doctor students and 2 people trained in CPR as they are part of the medical field (eg., nurses).

Song name	**Duration (min)**	**Language**	**Fit for CPR ?**	**Comments from experts / Natural health service of respective countries and other sources**
Bee Gees - Stayin Alive	4.1585	English	Yes	Highly recommended song for performing CPR[2][3][4]
Enrique Iglesias - Cuando Me Enamoro	3.350	Spanish	No	Not recommended for CPR as bpm is 60 (<100)
Daddy Yankee - Gasolina	3.214	Spanish	Yes	Highly recommended song for performing CPR[4]
iKON - Love scenario	3.5286	Korean	No	Not recommended because of varying BPM for half-time 59 (<100) or double-time at 236 (>200)
I'm Yours - Jason Mraz	5.044	English	No	Not recommended for CPR as bpm is less than 100

CPR Beat detection algorithm:

The beat detection algorithm consists of the following steps:

1. Converting input audio to wav format

The basic step used in this algorithm is to convert the audio format of the uploaded audio to a standard format for the ease of processing. **Ffmpeg package in python** will be used to convert these audio files to WAV format.

The common format is decided to which audio files will be converted to is WAV format as :

- MP3 formats are very space consuming because of its preference on sound quality.
- Most of the librosa apis used in this project work on top of WAV format

Note: It was observed that the Ffmpeg library used to convert audio files to wav format worked very fast at an average response time of 0.00045 seconds (0.45 ms).

Hence, using this library for audio file format conversion to wav is highly recommended.

2. The CPR beat detection

In order to perform effective hands only cpr the audio clip must follow the following conditions:

- beats per minute(bpm) of audio must be between 100 to 200 [5].
- Audio length must be of 2mins as CPR must be performed continuously for 2 mins before checking for sign of consciousness or breathing

Librosa algorithms

The following librosa apis were used to check if an audio clip satisfies the above criteria:

➔ **Librosa load**: To convert the wav file into digital signal for further analysis. The following is the time taken to do the above for the songs under observation:

Song name	**Song duration(sec)**	**Time taken (sec)**
Bee Gees - Stayin Alive	4.1585	11.1878
Enrique Iglesias - Cuando Me Enamoro	3.350	9.4059
Daddy Yankee - Gasolina	3.214	8.816
iKON - Love scenario	3.5286	13.202
I'm Yours - Jason Mraz	5.044	10.255

*Note: Since **librosa load** takes almost **11sec** for a 4 min song clip, it is used only once in the whole algorithm to convert audio to digital signal. This is because when **used to clip the window of audio and convert it to a digital signal** and run it through the algorithms, it took **~804.5665 sec** on an average to do so for the below discussed strategies.*

➔ **Librosa beat**: To check the tempo (bpm) of an audio. The librosa beat detection api for the above audios took an average of **~1.257ms** for each time it's used. From the response time point of view, this api is totally recommended for usage.

➔ **Librosa get_duration**: To get the duration of audio and predict maximum iterations it might take for each strategy tested.
Note: librosa getduration doesn't put the contents to be tested into memory for retrieval. Hence, it is very fast in querying the duration of long files.

The following are the different strategies used to find the best algorithm that satisfies the objectives of this project:

Note: For all of the following strategies, if audio length is less than 2 mins then algorithm is not run for that audio

CPR beat detection strategies

*Note: All strategies have a common **reject fast condition** that is to reject running any audio that is less than 2 mins through the algorithms, thereby saving a lot of computing time and memory and giving more time for other priority jobs to run.*

2.1. Finding the average bpm of the whole song

The librosa beat detection library is used for the same. Running the above songs through this strategy provided the following results:

Song name	**Time taken (sec)**	**Tempo (bpm)**	**Fit for CPR?**
Bee Gees - Stayin Alive	12.36	104	Yes
Enrique Iglesias - Cuando Me Enamoro	10.3	84	No
Daddy Yankee - Gasolina	10.24	96	No
iKON - Love scenario	14.274	118	Yes
I'm Yours - Jason Mraz	8.577	76	No

Observations made:

- Since few **portions of the song** have **slower** beats (melody music) and part of the song might break into a rap**(faster beats)**, this may **bring down or bring up** the **average bpm** of the whole song, leading to incorrect results. For example,
    - This can be observed from the results of the song “Daddy Yankee - Gasolina” which came out unfit for CPR as there’s a small portion of the song that is non CPRable.
    - **Worser case is when a non-CPR able song** “iKON - Love scenario” **comes up as CPRable** because of this anomaly

- The librosa load algorithm takes almost 10x more time than the time taken by the strategy itself.
    - average time taken by librosa to generate signal from wav: 12.0579ms
    - average time taken by ffmpeg to convert file format to wav: 0.0005ms
    - Hence the **signal generation is the most time consuming activity** and this api needs to be improved if such projects or applications need to be built on top of it.

2.2. Finding the bpm for every 2 mins with offset of 1 second

Effective hands only CPR needs to be performed every 2 mins with breath check until the patient starts breathing or help arrives. A minimum of 2mins of song length with required bpm (100 -120) is required. Hence, finds the bpm of the audio for offset of 1sec and duration of 2mins, until reaching the last 2 mins of the audio file.

Song name	Time taken (sec)	number of offsets	Tempo (bpm)	Fit for CPR?	Comments
Bee Gees - Stayin Alive	84.604	129	104	Yes	Interval(sec) - 0 - 249
Enrique Iglesias - Cuando Me Enamoro	57.193	81	85-125	No	Varying bpm observed
Daddy Yankee - Gasolina	52.26	72	96	No	Same ceiled bpm value observed
iKON - Love scenario	61.95	91	118	Yes	Interval(sec): 0 - 211
I'm Yours - Jason Mraz	116.85881	182	75-160	Yes	Interval(sec): 179 - 301

Observations made:

- This algorithm definitely caught the varying beats for most of the songs as shown below:
  Note: the x axis represents start time of beat analysis, y axis represents the tempo

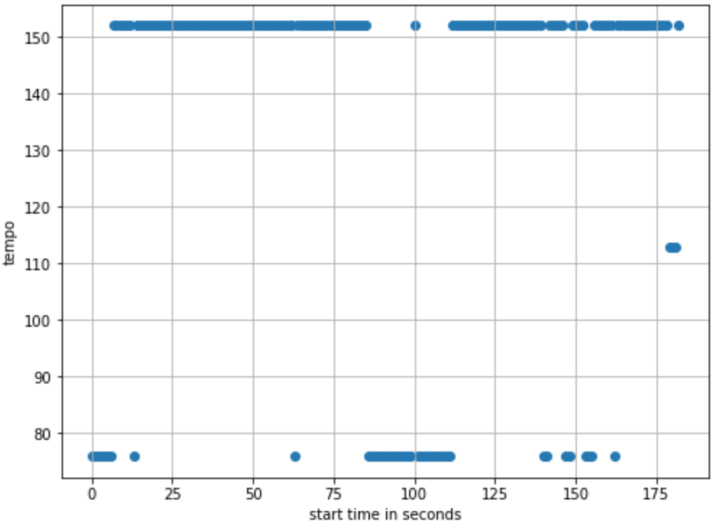


Figure 1. Observation of how beats found vary with sliding window of 1 second for the song "I'm Yours - Jason Mraz"

- The average amount of **time taken** by the **algorithm** is ~74.573 sec, which is **very high**.

- Some songs having repeating fast choruses does average out the overall average beat of the song like that observed in the song "Daddy Yankee - Gasolina". Hence, issues from the previous strategy are not entirely fixed with a sliding window of 1 sec.

If a segment of the song doesn't have needed bpm, then it is not worth the analysis. Hence, from this point forward different and bigger sliding window offsets were experimented.

2.3. Finding the bpm for every 2 mins with offset of 30 seconds on on poor CPR beat detection

Song name	Time taken (sec)	number of offsets	Tempo (bpm)	Fit for CPR?	Comments
Bee Gees - Stayin Alive	90.276	129	104	Yes	Interval(sec) - 0 - 249 Worst case scenario
Enrique Iglesias - Cuando Me Enamoro	11.51	3	84	No	Best case scenario
Daddy Yankee - Gasolina	11.639	3	96	No	Best case scenario
iKON - Love scenario	56.098	91	118	Yes	Interval(sec): 0 - 211 Worst case scenario
I'm Yours - Jason Mraz	14.58	8	75-160	Yes	Interval(sec): 180 - 301

Observations made:

- Produces the same results as that of the previous strategy with **much less iterations** and **total time taken**.
- Performs the **best** in case of **fast rejection** scenarios and the **worst when the whole song is CPR-able.** For example, best case scenario audio - “Enrique Iglesias - Cuando Me Enamoro” and worst case scenario audio - “Bee Gees - Stayin Alive”

2.4. Finding the bpm for every 2 mins with offset of 1 minute on on poor CPR beat detection

Song name	Time taken (sec)	number of offsets	Tempo (bpm)	Fit for CPR?	Comments
Bee Gees - Stayin Alive	86.2	129	104	Yes	Interval(sec) - 0 - 249 Worst case scenario
Enrique Iglesias - Cuando Me Enamoro	10.83	3	84	No	Best case scenario
Daddy Yankee - Gasolina	10.713	2	96	No	Best case scenario
iKON - Love scenario	59.005	91	118	Yes	Interval(sec): 0 - 211 Worst case scenario
I'm Yours - Jason Mraz	11.98	4	110-160	Yes	Interval(sec): 180 - 301

Observations made:

- Produces the same results as that of the previous strategy with **much less iterations** and **total time taken** for the **best case scenarios**

- **Worst case scenarios** still **perform the same**

2.5. Finding the bpm for every 2 mins with offset of 2 minute on on poor CPR beat detection

Song name	**Time taken (sec)**	**number of offsets**	**Tempo (bpm)**	**Fit for CPR?**	**Comments**
Bee Gees - Stayin Alive	88.019	129	104	Yes	Interval(sec) - 0 - 249 Worst case scenario
Enrique Iglesias - Cuando Me Enamoro	10.39	1	84	No	Best case scenario
Daddy Yankee - Gasolina	10.41	1	96	No	Best case scenario
iKON - Love scenario	59.41	91	118	Yes	Interval(sec): 0 - 211 Worst case scenario
I'm Yours - Jason Mraz	10.074	3	75-160	No	Strategy couldn't find the CPR interval

Observations made:

- **Since a song is itself 3-5 mins long, a sliding window having an offset of 2mins seems to lead to poor CPR interval detection compared to previous strategies.**

In order to encourage faster CPR interval detection, the last strategy follows a binary search technique where every 2 mins is checked for CPR beat

2.6. Finding the bpm of for every 2 mins as offset

Song name	**Time taken (sec)**	**number of offsets**	**Tempo (bpm)**	**Fit for CPR?**	**Comments**
Bee Gees - Stayin Alive	13.13	2	104	No	Different results from other strategies. Method concludes audio not CPR suitable
Enrique Iglesias - Cuando Me Enamoro	10.8	1	84	No	
Daddy Yankee - Gasolina	10.27	1	96	No	
iKON - Love scenario	6.505	1	118	No	
I'm Yours - Jason Mraz	9.305	2	76	No	Least time taken so far

Observations made:

- The poor CPR interval detection due to the 2 min sliding window is observed to have much worse effects compared to the previous strategy.
- **None of the CPR-able intervals were detected through this strategy.**

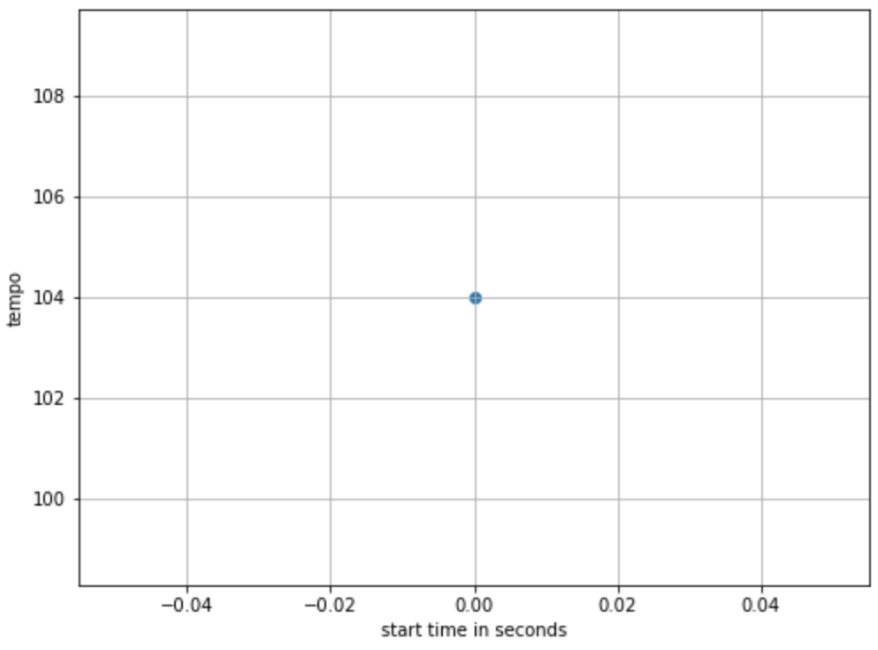
Evaluations

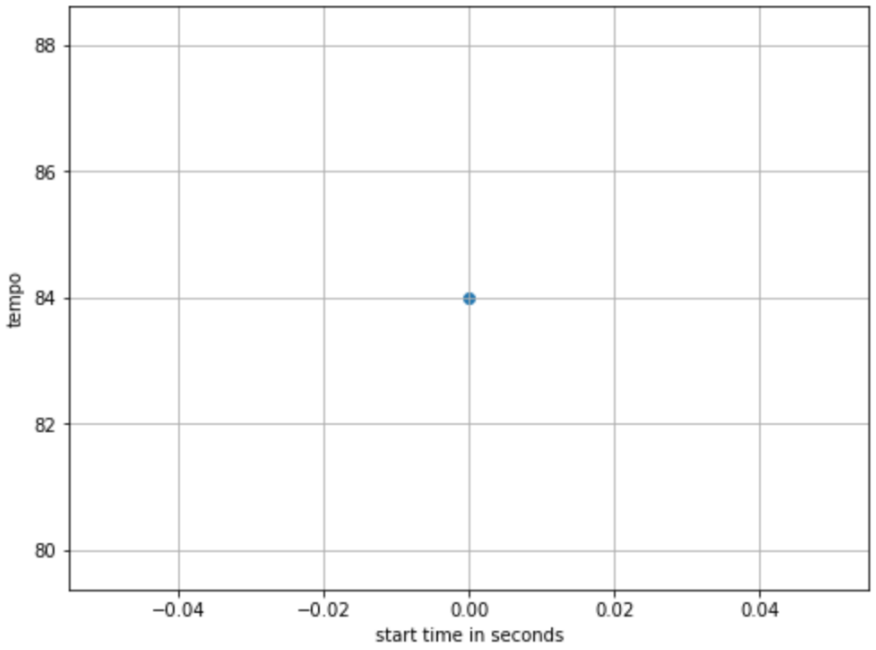
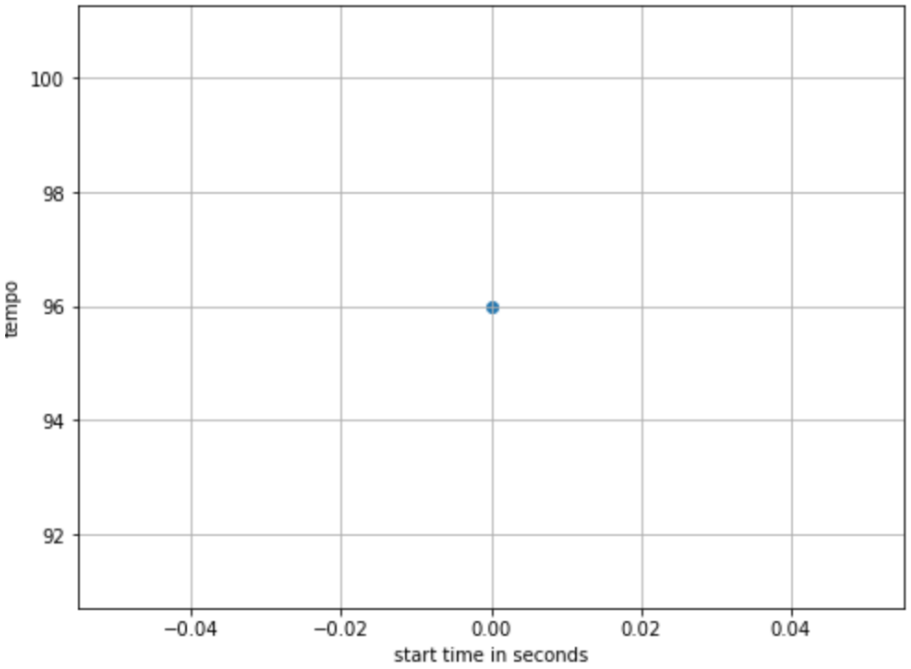
1. Objective evaluation

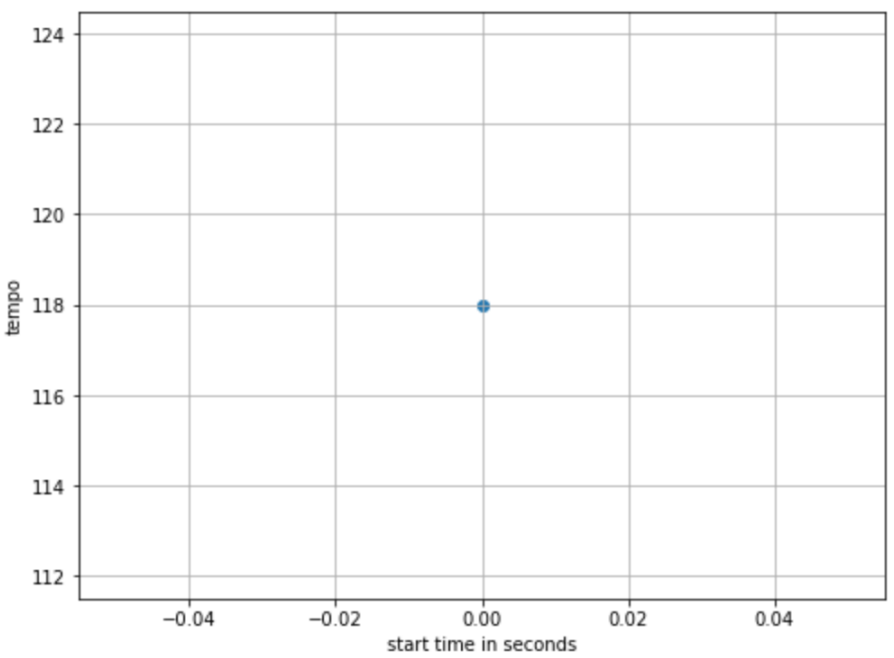
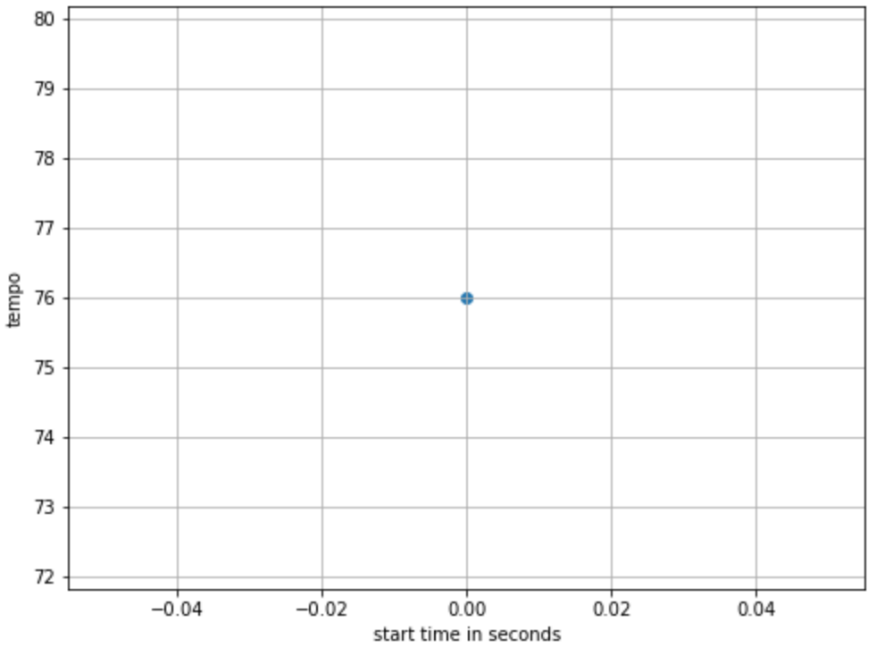
   The **worst strategy** to find CPR-able interval in an audio was found to be the binary search strategy of **finding the bpm of for every 2 mins as offset**

   The **best strategy** to find CPR-able intervals in an audio is - **Finding the bpm** for every 2 mins with an offset **of 1 minute on poor CPR beat detection**. It produced the most optimal results at better time and for very iterations.

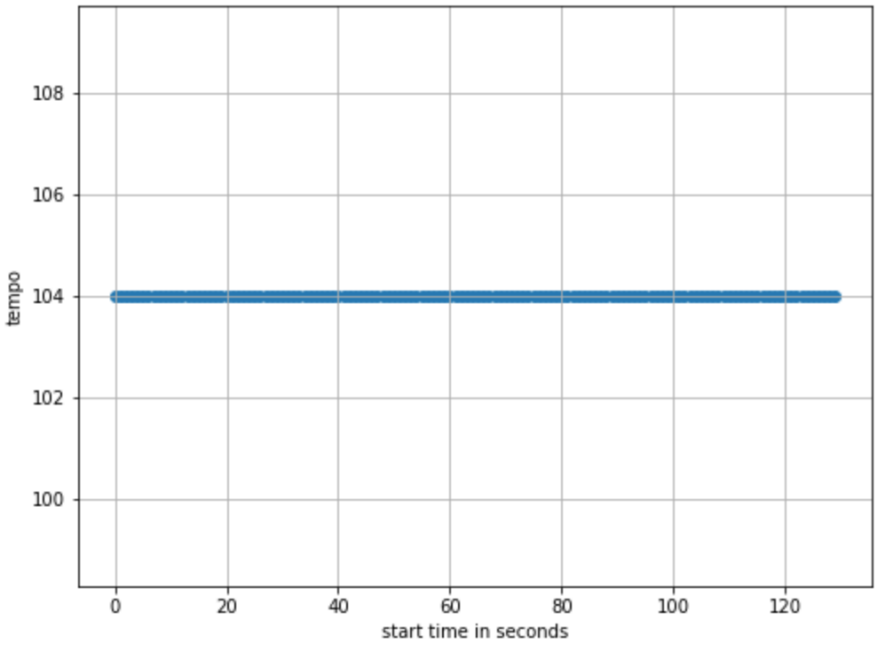
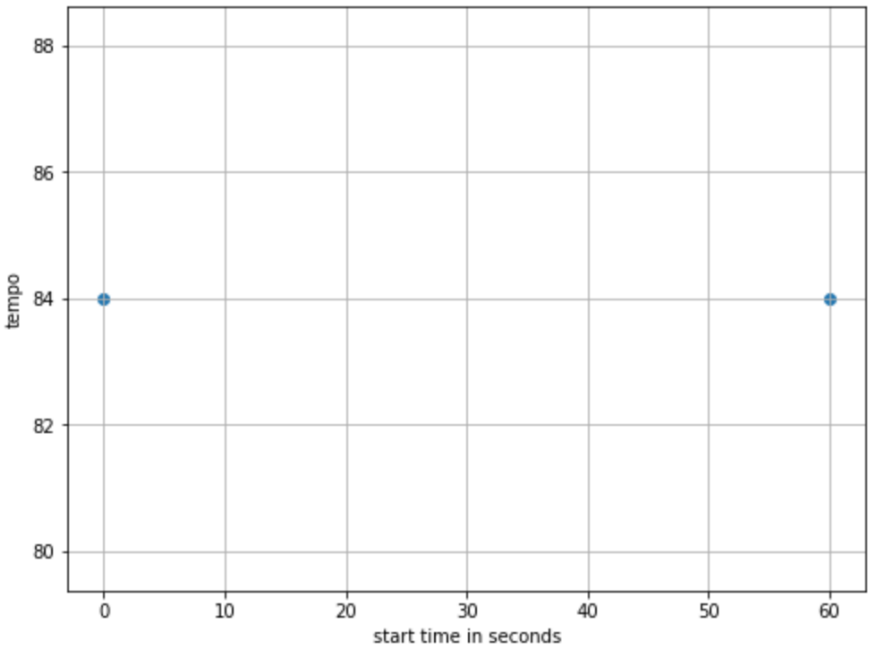
   The following will show the difference between optimal CPR beat start times detected for the worst and best strategy for each song:

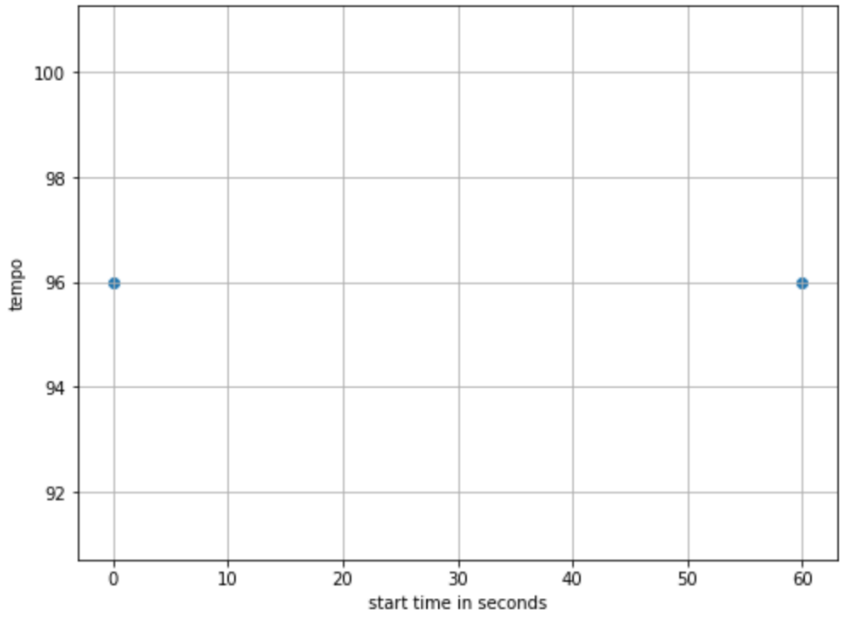
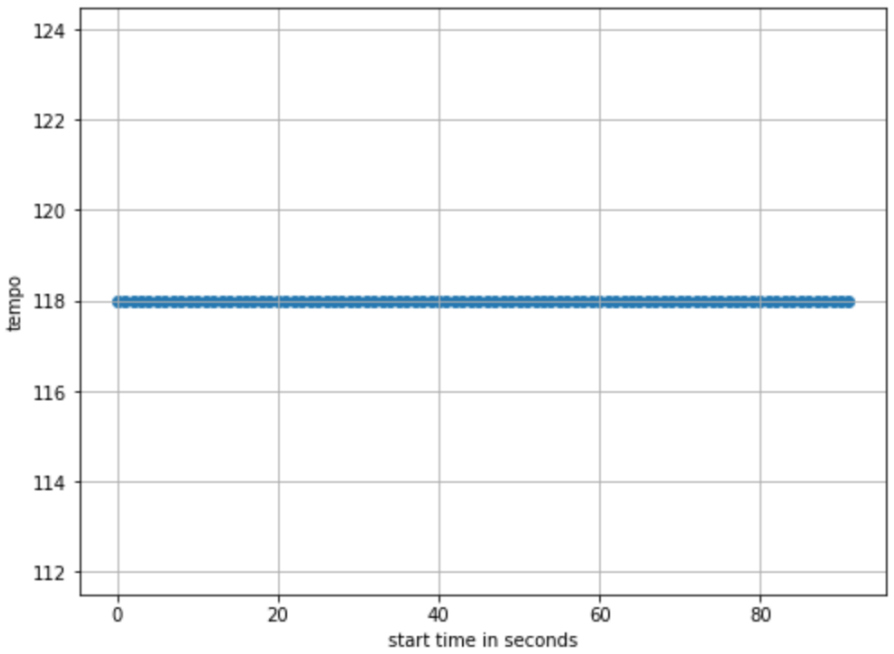
Song name	Worst strategy (strategy 6)
Bee Gees - Stayin Alive	

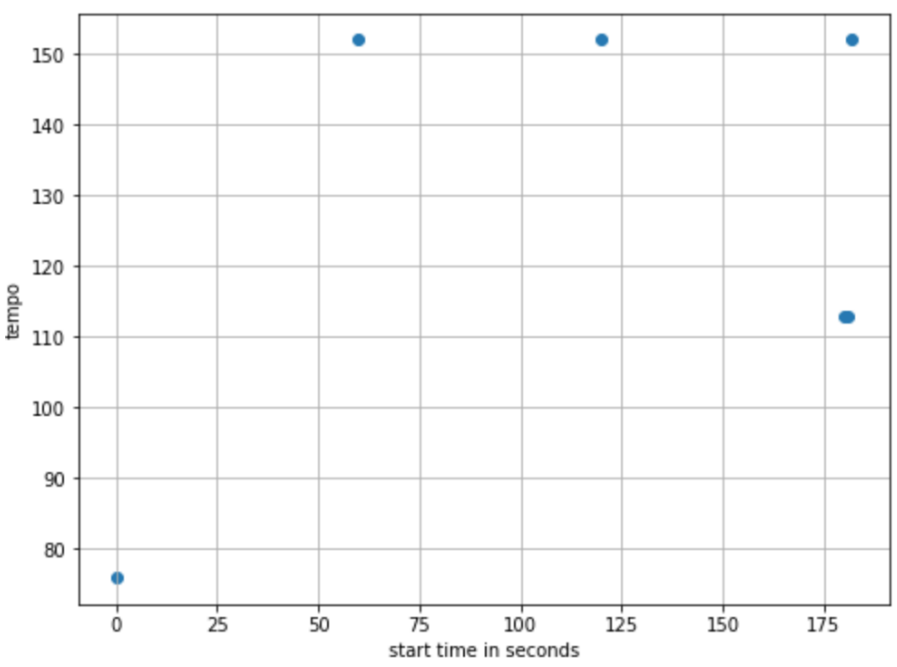
Enrique Iglesias - Cuando Me Enamoro	
Daddy Yankee - Gasolina	

iKON - Love scenario	
I'm Yours - Jason Mraz	

Quick Note : ***Plotting of these graphs was faster when using normal arrays instead of using ndarrays*** *which was how we did for our practical assignments. This is another observation done when doing these evaluations that some developers might find it to be useful.*

Song name	Best strategy (strategy 4)
Bee Gees - Stayin Alive	
Enrique Iglesias - Cuando Me Enamoro	

Daddy Yankee - Gasolina	
iKON - Love scenario	

I'm Yours - Jason Mraz	

2. Subjective evaluation

The subjects for this evaluation includes 42 people from medical backgrounds like doctors, nurses, CPR trainers, etc and those who are not CPR trained. The aim of the evaluation is to find if findings from the best strategy (strategy 4) matches (or not) with the results from subjects with and without medical background along with the data collected from experts which was used as initial reference point for correctness of the algorithm.
This is the form containing questions for the subjective analysis
https://golisten.ucd.ie/task/ab-test/644ab391b72736ea6b003e33
The Go listen application was used for the same purpose.
Note: This application doesn't work for IOS phone devices and hence a lot of Iphone users found it difficult to take the test using their phones.

The form collects the following data

- The subject's profession
- The subject's hearing ability
- What language music do they usually listen to
- **A reference to a sample CPR beat audio**
    - **This will be useful for people from non medical background to have a reference and answer the questions.**
    - **If positive results in detection is seen with non CPR trained people then the purpose of the project is proven**
- If the subject same 5 audios (same as the ones used so far) CPR able

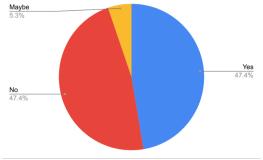
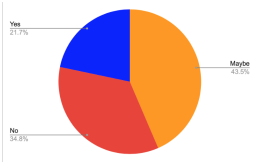
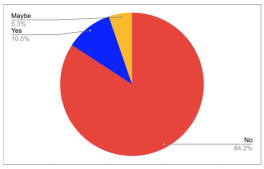
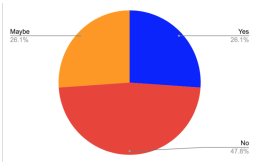
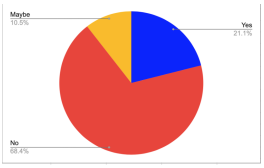
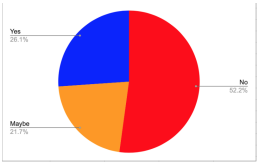
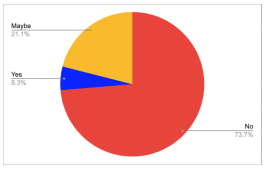
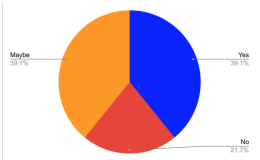
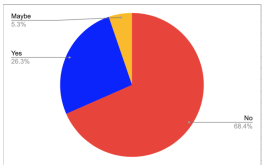
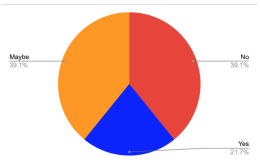
The following observations are made on top of :

- 19 medical professional/CPR trainers - 5 doctors, 5 medical professionals and was trained during my studies or work, 9 officially CPR trained people under private

organisation

- 23 non CPR trained general people

The blue indicates the Yes, Red the no's and orange for maybe

Song name	**Fit for CPR as per medical professional/CPR trainers ?**	**Fit for CPR as per non CPR trained general public?**	**Fit for CPR by experts (initial reference)**	**Fit for CPR as per best performing beat detection algorithm?**
Bee Gees - Stayin Alive			Yes	Yes
Enrique Iglesias - Cuando Me Enamoro			No	No
Daddy Yankee - Gasolina			Yes	No
iKON - Love scenario			No	Yes
I'm Yours - Jason Mraz			No	Yes

Observations made:

- Many **"maybe" answers** from the general **public who are not CPR trained but listened to the reference CPR** audio as compared to that of trained professionals. This can be seen in 2 ways:
    - Existence of overconfidence bias from the trained professionals. M. Barber and Odean[6] introduced the concept of overconfidence in the paper written in "The Quarterly Journal of Economics" (February, 2001).
    - We always give ourselves the benefit of the doubt even though we know it's right. Hence, for this reason 50% of the "maybe" answers are considered as

yesses for the upcoming investigations.

- **The results of the beat detection algorithm matches better with answers from the general public compared to that from experts and that of professionally trained individuals.**
    - From looking at the music preference between CPR trained and untrained public, the English and other language preference level is almost the same. Only for Indian languages it differs and this test doesn't contain any indian language songs, hence language of preference isn't the reason for this anomaly.

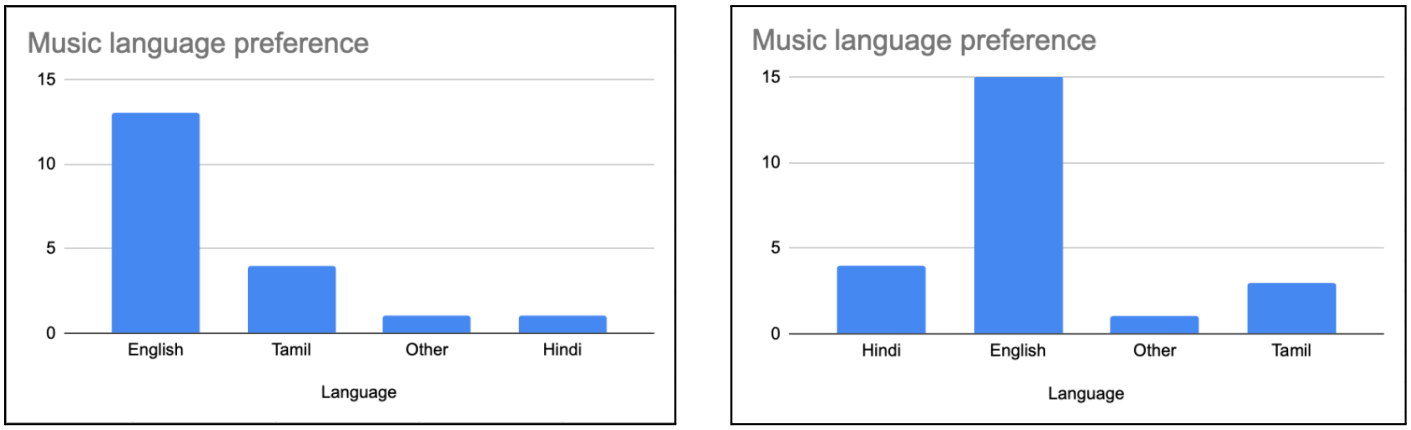


Figure 2. Music prefer of CPR trained professionals (left) and non CPR trained general public (right)

    - **The songs were evaluated incorrectly mostly by nurses and doctors.** After contacting the experts and sending out follow up FAQs, I was able to understand the possible reason behind this puzzlement.
        - Doctors and other medical professionals are not CPR trained using music, but by time beats. Hence, half of the professionals who answered the questions **thought that** I had asked **if they would play this music when performing CPR**.
        - **The subjective evaluation survey questions should have been better framed to fit audience of all types**

- **The subjects chosen are biassed as** 70% of them listen to english music
    - Hence, languages closest to English, i.e, Spanish were identified as expected.
    - However, CPR-able Korean audio ("iKON - Love scenario") wasn't recognised, thereby proving the hypothesis on which this project was set.

Topics covered from lecture:

- Digital representation of sound
- Basic audio processing - segmentation and window sizing
- With the subjective tests a little bit of psychoacoustic tests is covered in the sense how is different sound being perceived by different people

- Speech intelligibility with reference as CPR beats, subjective and objective testing mechanism
- Topics covered from practicals:
    - Librosa apis for loading signal, beat detection, finding signal duration etc
    - Python apis - plotting

Reflections:

- Librosa load definitely needs to be improved by the developing community, more so that since it's open source more people can contribute to its development. Digital representation of signals is the basis for most of the audio analysis and this project analysis proves that this process is the most time consuming activity.

- Along with above suggested strategies, one more suggestion would be to quit from the algorithm post finding the first 2 mins that satisfies the CPR audio conditions. Since this contradicted with the objective of the project which is to find the longest duration more than or equal to the needed duration of 2mins, this strategy wasn't implemented and tested

- Since a song is itself 3-5 mins long, a sliding window or having an offset of 2mins leads to poor CPR interval detection.

- The **worst strategy** to find CPR-able interval in an audio was found to be the binary search strategy of **finding the bpm of for every 2 mins as offset**

- The **best strategy** to find CPR-able intervals in an audio is - **Finding the bpm** for every 2 mins with an offset **of 1 minute on poor CPR beat detection**. It produced the most optimal results at better time and for very iterations.

- **The recorded approximate response values for each ot the strategies will change with change in audio size**

- There definitely needs to be better applications other than golisten that can work in all types of devices and can do audio based AB tests. For example, IOS mobile users couldn't not open the tests created in the golisten app using their iphones as this application is not supported for IOS.

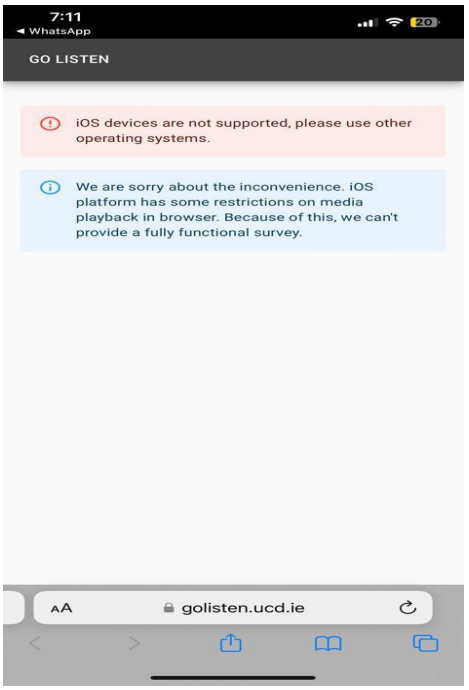


Figure 3. Screenshot showing golisten application not opening in IOS devices along with the error message

- **General public when provided with a reference to CPR beat were able to identify audio which are CPR-able. Hence, proving the value of this project especially for the general public.**

- **The subjective evaluation survey questions should have been better framed to fit audience of all types**

- **The subjects chosen are biassed as** 70% of them listen to English music and **hence weren't able to identify CPR-able audio in another language**, Korean.

- Practical benefits of the application seen from user's point of the view
    - Since the project provides an interval of 2+ mins to which the user of the application can perform CPR to, it allows them to have a mental frame for the 2 mins instead of being in the situation to check their clock or keeping an alarm, etc, during the state of panic.

References:

- [1] Pellegrino JL, Vance J, Asselin N. The Value of Songs for Teaching and Learning Cardiopulmonary Resuscitation (CPR) Competencies: A Systematic Review. Cureus. 2021 May 16;13(5):e15053. doi: 10.7759/cureus.15053. PMID: 34141503; PMCID: PMC8204400

- [2] Paul Martin 26 August 2022, *proCPR.org website*, accessed 20 April 2023, <https://www.procpr.org/blog/training/cpr-stayin-alive-song>

- [3] Anonymous 5 March 2022, *youtube Fire Department Chronicles channel,* accessed 20 April 2023, <*https://www.youtube.com/watch?v=0H_XUaxrna8*>

- [4] Paul Martin 1 July 2022, *proCPR.org website*, accessed 20 April 2023, <https://www.procpr.org/blog/training/cpr-chest-compression-rate>

- [5] Anonymous 15 March 2022, nhs website, accessed 16 April 2023, <https://www.nhs.uk/conditions/first-aid/cpr/>

- [6] Brad M. Barber AND Terrance Odean, Boys will be boys: gender, overconfidence, and common stock investment, The Quarterly Journal of Economics, February 2001