Chord Recognition using Machine Learning

February 5,2022

Aditi Kanzarkar
Computer science
PVG's College of Engineering and Technology & G. K Pate (Wani) Institute of Management
Pune, Maharashtra
aditi.kanzarkar@gmail.com

Kashmira Dhere
Computer Science
PVG's College of Engineering and Technology & G. K Pate (Wani) Institute of Management
Pune, Maharashtra
kashmiras143@gmail.com

Vicky Mali
Computer Science
PVG's College of Engineering and Technology & G. K Pate (Wani) Institute of Management
Pune, Maharashtra
malivicky1118@gmail.com

Jozia Rashid Khan
Computer Science
PVG's College of Engineering and Technology & G. K Pate (Wani) Institute of Management
Pune, Maharashtra
khanjozia3@gmail.com

Prof.Aditya G. Dongre
Department of Computer Engineering, Pune Vidhyarthi Griha's College of Engineering and Technology
Savitribai Phule Pune University
agd_comp@pvgcoet.ac.in

ABSTRACT

In this paper, we consider the difficult issue of music acknowledgment and present a successful profound learning-based strategy involving Feed forward neural network for chord acknowledgment. Profound learning has become well known for some, music handling errands with Neural Networks (NNs) frequently applied. As the PCP has brilliant element of, conservative portrayal of the recurrence content of a sound, in this review, we stretch out the PCP vector plan to address the constraint. Our proposed strategy essentially comprises of two significant stages. To start with, we present novel channels and apply then to PCP vector to change the vector into the 12 notes of a standard chromatic scale. The subsequent advance is to proficiently learning highlight on the changed framework utilizing Feed Forward neural organization. We propose a teachable, information driven approach that naturally learns includes and its classifier all the while.

KEYWORDS - PCP, Machine Learning, Feed-Forward Neural Network, Chroma Feature

I. INTRODUCTION

Harmony acknowledgment alludes to a course of endeavouring to recognize a harmony from a piece of sound. The record of harmonies has been completed physically which is extremely tedious, tedious and including the information on the melodic. Hence, harmony acknowledgment is a troublesome assignment to make due. It has required the verifiable melodic information as well as worried about the blunders made by human. Melodic harmonies can be characterized by any consonant arrangement of at least two notes for the most part played together in a bunch of concurrent tone. Subsequently programmed apparatuses play a part to play. Music Information Retrieval (MIR) is an interdisciplinary science whose objective is to broaden data recovery into non text based just regions. The point of MIR is to depict various viewpoints connected with the substance of music. A few uses of MIR incorporate music record, music characterization, playlist age, and music acknowledgment. Here our undertaking profoundly centres around utilization of MIR that is music acknowledgment Several strategies are accessible to perceive music yet the most likely most popular to performers is that of harmonies. A harmony can be characterized as a bunch of concurrent tones.

This definition may be suitable for a human however, according to a characterization perspective, it

seems, by all accounts, to be improper in light of the fact that there are numerous varieties (because of the instruments, clamour, recording conditions, and so on) in any event, for an extraordinary harmony and this make issue while perceiving harmonies for obscure music. As the PCP has incredible element of, smaller portrayal of the recurrence content of a sound communicated as the general extent of energy as for the 12 notes of a standard chromatic scale and as the vast majority of the concentrated on PCP highlights are touchy to the sounds relying upon the instrument, and different boundaries like dynamic, assault and support ,PCP turns into the principal part for making harmony acknowledgment framework and becomes appropriate applicant as it has a little aversion to instrument change

As the reason for this undertaking is to make the exact harmony acknowledgment framework utilizing AI as opposed to utilizing successive matching which certain individuals have done in before time and to give the precise result as harmony sheet for obscure tune, we will actually want to see further that utilization of AI models and PCP assists with making great harmony acknowledgment framework without manual endeavours.

II. MOTIVATION

Envision you hear some intriguing tune and might want to figure out how to play it on guitar or piano. Commonly you want text (verses) and tune to sing and chords (different tones without a moment's delay) for backup. Here and there we're fortunate since somebody interpreted them on paper (or website page) previously, some of the time not. Text and tune are typically simple to pick by ear, however chords can be some of the time interesting regardless of whether you're prepared well. Accordingly, it would be great if a machine would assist us with perceiving what chords are there in the sound. To gain proficiency with another tune on instrument you really want verses, song and chords for backup.

We can help verses through web and can get tune effectively simply by hearing the melody yet issue comes when we can't get the ideal chords for the tune to play that tune on specific instrument. By and large Chord sheets (for example sheet include chords) are composed by thoroughly prepared artists and they are distributed for others to use, on web however this is completely manual undertaking which is finished by specific individual. As this undertaking is manual there are issues appended with it generally for individuals who are amateurs in field of music and have want to master utilizing web. Some issues frequently happen which are as follow,

1. Manually written chord sheets are not present on web for your favourite songs

2. You are unable to recognize which instruments are present in song for covering that song.

3. The song you want to play doesn’t have free availability of chord sheet

so, the whole point of this project is to solve above issues using ML.

III. LITERATURE SURVEY

a) Related Work

Chroma features, also known as Pitch Class Profiles (PCP), have been used as front end to chord and songs recognition systems from audio recorded queries.

Specifically, it has been shown that chroma highlights are appropriate forcover tunes identification frameworks. PCP highlights are great mid-level elements which supportive of vide a more solid and direct portrayal of tunes than song.

The first PCP was presented by Fujishima in1999. In this PCP, the powers of all recurrence receptacles inside the limits of a semitone are summarized and the semitones in octave distance are amounted to pitch classes, bringing about a 12-canister PCP vector. Varieties of this vector incorporate 24-container and 36-canister vectors, bringing about additional exact highlights. Fujishima utilized his PCP vector to perform design matching utilizing double harmony type formats (i.e., optimal PCP portrayals as displayed in Figure 2).

Lee presented another element vector called the Enhanced Pitch Class Profile (EPCP) for programmed harmony acknowledgment from the crude sound. To this end, he first acquired the Harmonic Product Spectrum (HPS) from the consistent Q change (CQT) of the information sign and afterward he applied a calculation to that HPS for processing a 12-layered improved pitch class profile. The CQT has mathematically divided focus frequencies which can be dimensioned so they relate to melodic notes. It is accordingly an intriguing pre-handling venture for music registering.

Gomez and Herrera proposed a framework that naturally separates, from sound accounts, apparent metadata like harmony, key, scale and rhythm data. In their work, they registered a vector of low-level momentary elements: the HPCP (Harmonic Pitch Class Profile) vector. It depends on the power of each pitch planned to a solitary octave, which compares to Fujishima's PCP.

Harte additionally proposed a strategy involving the CQT for harmony acknowledgment. What's more, he added a tuning calculation which can manage varieties in instrument tuning.

Sheh and Ellis proposed a measurable learning technique for harmony division and acknowledgment, where they utilized Hidden Markov Models (HMMs) prepared by the Expectation Maximization (EM) calculation, and treated harmonies names as covered up values inside the EM system. The greater part of the previously mentioned work on harmonies acknowledgment doesn't utilize AI methods, but instead involves signal handling strategies to acquire the most ideal 12-receptacle PCP vectors, and afterward perform design coordinating. Nonetheless, it is very difficult to get an ideal 12-canister PCP vector which features just the principal notes of a harmony. For sure, each instrument brings new sounds, and the dynamic of the performer, among different boundaries, adds clamor to the PCP. Thus, we propose a framework in view of AI strategies, whose objective is to gain proficiency with an appropriate model typifying this large number of boundaries.

In any case, no genuine marked harmonies information base is by all accounts openly accessible (as far as anyone is concerned) to fabricate such a model. In this work, we propose an information base, and we consider the utilization of genuine harmonies tests to prepare a more precise harmony acknowledgment framework. Since we want to involve our framework for music acknowledgment, we really want quick calculations, which are important to manage immense information bases of melodies. Subsequently, we decided to utilize the first PCP vector since it is quick and includes not many pre-handling steps.

b) Current Market Survey

1. Chord Tracker - Have you at any point attempted to sort out what the harmonies are to your main tunes? Yamaha's new Chord Tracker application does the difficult job for you, and substantially more! The Yamaha Chord Tracker application helps you practice and perform tunes by investigating the sound melody put away in your gadget and afterward shows the harmony images for you.

2. MyChord - Do not meander the Internet searching for harmony scores any longer. See the harmonies showed alongside the music and play with instruments like guitar and piano.

IV. CHROMA FEATURES

a) Principle

The most commonly used descriptor for chord identification has been the Pitch Class Profile (PCP). The equation for the full PCP algorithm which is described by Fujishima is as follows:

$$PCP(p) = \sum_{ls.t.M(l)=p}^{l} \|X(l)\|^2$$

$$M(l) = round\left[12\log_2\left(\frac{f_s t}{f_{ref}}\right)\right] mod\ 12$$

where fref is the reference recurrence you are attempting to coordinate, N the quantity of canisters in the FFT of the information signal or normally known as FFT size, note that a bigger N (more examples, longer time) would mean higher goal of the FFT. Furthermore, fs is the examining recurrence of the sound record you are handling. Regularly, this is 44100 Hz.

This is normally a bunch of 12 frequencies, addressing the key frequencies of the 12 semitones of a traditional console. For your reference, they are [16.35, 17.32, 18.35, 19.45, 20.60, 21.83, 23.12, 24.50, 25.96, 27.50, 29.14, 30.87], beginning from C0 to B0.

To comprehend this condition better, the aggregate term of fsl/N is comparable to changing over the FFT canisters into the genuine recurrence that container addresses." l "here is only the receptacle counts for the FFT yield. Along these lines, for each container of the FFT, the partner recurrence can be determined with this term.

Since each octave is really a twofold in recurrence to the past octave (in the event that C0 is 16.35Hz, C1 = 2*C0 = 32.70 C2 = 2*C1). Along these lines, to switch this, we take the log2 of the determined recurrence from the past advance.

Presently to the principal condition, note the summation if for all "l", with the end goal that M(l)=p. This implies for all the different pitch classes p (0-11), we should only involve receptacles in the FFT along with the end goal that M(l)=the pitch classes. To observe such circumstances, we clear the fref for all frequencies in the FFT so we observe the frequencies that contributes the most to the singular pitch classes.

Note that the M(l) condition shouldn't give you a negative number. The justification for this is the FFT makes both positive and negative frequencies. In any case, for every genuine sign, (for example, sound signals), the negative frequencies are only an impression of the positive frequencies. You just truly need to manage the positive frequencies, which dwells in the primary portion of the result from the FFT. Or on the other hand you can utilize a rFFT (like the one in python - NumPy) calculation that explicitly manages genuine signals, and will just re-visitation of you the positive frequencies.

b) Normalization of PCP

To look at PCP vectors, normalizing them is essential. To be sure, a harmony can be played stronger or gentler and in this way, energy circulation can fluctuate starting with one preliminary then onto the next. To standardize a PCP vector, we partition the energy of each container by the absolute energy of the first PCP, that is to say,

$$PCP(P) = \frac{PCP^*(p)}{\sum_{j=0}^{11} PCP*(j)}$$

where p is the index of the bin is to be normalized.

V. CHORD RECOGNITION SYSTEM

a) Feed Forward Neural Network

i) Definition

A feedforward neural network is a naturally inspired request computation. It includes a (possibly tremendous) number of fundamental neuron-like dealing with units, composed in layers. Every unit in a layer is related with all of the units in the past layer. These affiliations are not all same: each affiliation could have a substitute strength or weight. The heaps on these affiliations encode the data on an association. Regularly the units in a cerebrum network are furthermore called center points.

Data enters at the information sources and goes through the association, layer by layer, until it appears at the outcomes. During normal action, that is the place where it goes probably as a classifier, there is no analysis between layers. Hence, they are called feedforward cerebrum associations.

A Feed Forward Neural Network is conventionally found in its most direct design as a single layer perceptron. In this model, a movement of data sources enters the layer and are copied by the heaps. Every value is then added together to get a measure of the weighted data values. If how much the characteristics is over a specific edge, by and large set at nothing, the value conveyed is generally speaking 1, while expecting the all out falls underneath the limit, the result esteem is - 1. The single layer perceptron is a huge model of feed forward mind associations and is as a rule used in game plan tasks. Besides, single layer perceptron can consolidate parts of AI. Utilizing a property known as the delta rule, the brain organization can contrast the results of its hubs and the expected qualities, in this manner permitting the organization to change its loads through preparing to create more exact result values. This course of preparing and learning produces a type of an angle plunge. In diverse perceptron, the most common way of refreshing loads is almost closely resembling, but the interaction is characterized all the more explicitly as back-engendering. In such cases, each secret layer inside the organization is changed by the result values created by the last layer.

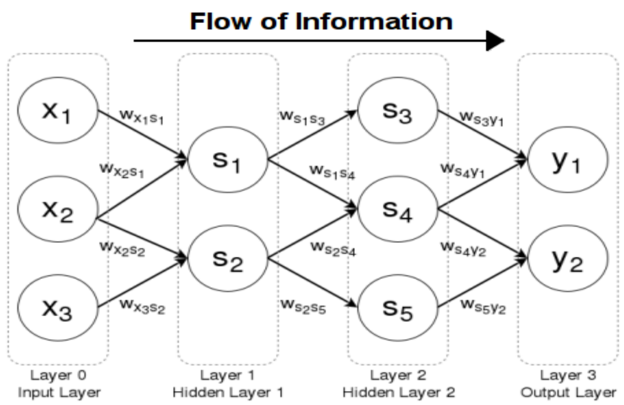

Figure 1: Feed Forward Neural network

The most straightforward sort of brain network is a solitary layer perceptron network, which comprises of a solitary layer of result hubs; the data sources are taken care of straightforwardly to the results by means of a progression of loads. The amount of the results of the loads and the data sources is determined in every hub, and assuming the worth is over some limit (regularly 0) the neuron flames and takes the initiated esteem (commonly 1); in any case, it takes the deactivated esteem (normally - 1). A solitary layer brain organization can process a persistent result rather than a stage work. A typical decision is the alleged calculated work:

$$f(x) = \frac{1}{1+e^{-x}} f(x) = \frac{1}{1+e^{-x}}$$

With this decision, the single-layer network is indistinguishable from the calculated relapse model, generally utilized in measurable demonstrating. The calculated capacity is one of the group of capacities called sigmoid capacities on the grounds that their S-molded charts look like the last letter lower instance of the Greek letter Sigma. It has a nonstop subsidiary, which permits it to be utilized in backpropagation. This capacity is likewise favored on the grounds that its subordinate is handily determined:

$$f'(x) = f(x)\,(1 - f(x))$$

(The way that f fulfills the differential condition above can be effectively be shown by utilizing the chain rule)

On the off chance that solitary layer brain network initiation work is modulo 1, this organization can tackle XOR issue with a solitary neuron.

$$f(x) = x1\ f(x) = x1$$

ii) Model Configuration

We have utilized keras ka sequential model. There are two thick layers in brain organization, in yield layer there are 10 neurons for 10 harmonies and 12 neurons in input layer for 12 notes of PCP. To upgrade our model, we have utilized misfortune capacities which are paired and straight out cross entropy, as the result classes are 10 consequently all out cross entropy misfortune work, which upholds multiclass model. Enactment capacities utilized are SoftMax which anticipate multinomial likelihood dissemination and depend. For the analyzer we have utilized Adam enhancer which is SGD strategy.

VI. DATABASE

Some portion of the hardships in AI procedures begin from the elaboration of a data set of tests, likewise called dataset, and harmony grouping is no special case. The essential prerequisite for a harmony recognizer utilizing AI procedures is that the dataset contains an adequate number of information to construct a model. The majority of the work portrayed in Section 2 on harmony acknowledgment doesn't utilize AI strategies which could make sense of why no harmonies information base is by all accounts openly accessible. Hence, we needed to make our own information base of harmonies. In our information base, we accumulated sound documents (kept in the WAV design, tested at f s = 44100 Hz, and quantized with 16 pieces), and the comparing precomputed PCP vectors. The PCP vectors were figured on windows containing each 16384 examples, which compare to 0,37 seconds. The window size was picked tentatively. We saw that windows containing just 4096 examples produce right outcomes, nonetheless, best outcomes for our application were accomplished utilizing a greater window size.

Since our last objective isn't to perceive every one of the current harmonies, however to foster a music acknowledgment framework, we can restrict harmonies to the most successive ones. Consequently, we picked a subset of 10 harmonies:

A, Am, Bm, C, D, Dm, E, Em, F and G.

In our data set, these harmonies are addressed by an identifier going separately from 0 to 9. Note that assuming different harmonies are additionally played in a melody, the fundamental harmonies can get the job done. Besides, numerous cutting edge melodies played in Western Europe depend on these 10 harmonies. In this manner, it is by all accounts a reasonable beginning stage to approve our acknowledgment strategy.

In useful terms, all PCP vectors are put away in an extraordinary record which is coordinated as follows. Each line comprises in a standardized PCP vector of twelve components and another component for the comparing harmony identifier. Coming up next is illustrative of one line of the dataset record

0.04, 0.09, 0.18, 0.05, 0.12, 0.04,

0.14, 0.04, 0.03, 0.18, 0.04, 0.05, 4

The last digit compares to the class (the D harmony in this model).

Then, we focus on the setting of the dataset. As we will approve our dataset and test it on examples procured in various settings, the information base was parted into two subsets. The principal subset contains an exceptionally huge measure of guitar harmony tests, though the subsequent subset contains a more modest arrangement of harmonies played with an alternate guitar and three different instruments. In this way, there are two different ways of utilizing the information base: we can either utilize cross-approval methods on the principal subset, or use it as a learning set while the subsequent subset is utilized as a test set. Subtleties follow.

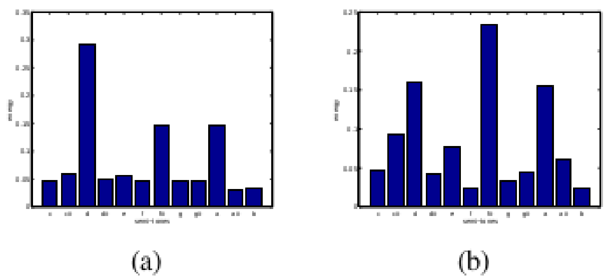

Figure 2: PCP representations of D chord recorded with the guitar in (a) an anechoic chamber, and (b) within a noisy room. Note: There are three major semi-tones are still visible in (b).

a) First subset

Chords of the dataset are created with an acoustic guitar, which is presumably the most well-known instrument in Western Europe to play harmonies. The securing conditions are the accompanying. The harmonies tests were kept in two distinct conditions. A big part of the harmonies were kept in an anechoic chamber with a wideband amplifier (01dB MCE320). The other half was kept in a boisterous climate, with a solitary live mouthpiece (Shure SM58). We felt that examples kept in the two conditions would reflect both the circumstances of experts playing their tunes in a studio and individuals playing music home. In Section 5, we determine that the framework performs best assuming it is prepared with a blend of commotion free and uproarious harmonies. It merits seeing that the harmonies were recorded utilizing a few playing styles (arpege, staccato, legato, and so forth.). Figure 3 shows the PCP portrayals of a D harmony kept in the anechoic bore and in the boisterous room with a similar guitar. As numerous genuine melodies are played in a boisterous climate, remembering uproarious harmonies for the database is pertinent. For each harmony, 100 examples were kept in the anechoic chamber, and 100 examples were kept in a loud room. For every climate, the examples were recorded utilizing four distinct guitars: one traditional guitar with nylon strings, and three acoustic guitars delivering three unique sounds. It is normal that the assortment of the dataset concerning guitars will upgrade the vigor of the framework and broaden its pertinence, as there are a wide range of guitars sounds accessible around the world. Taking everything into account, the principal subset is coordinated as follows. There are 2000 harmonies altogether. Every particular harmony is recorded multiple times, 100 in an anechoic chamber and 100 in a loud room. In both hundred parts, the harmonies are additionally isolated into four subsets of 25 harmonies, created with one of the four guitars.

b) Second subset

We additionally made a more modest information base containing harmonies recorded with a guitar and three different instruments, specifically a piano, a violin, and an accordion. That information base is expected to give a free test set and ought not be utilized to prepare the model. That information base contains 100 harmonies to for each instrument. These 100 harmonies are appropriated similarly among the ten harmonies referenced before. In this way, there are 10 examples for every harmony for each instrument. Set up Your Paper Before Styling

c) Pre-Processing

A .wav file is the contribution to the framework which needs preprocessing prior to passing it to the neural network model. In the preprocessing step we took the .wav input sound records and executed the pitch class profile calculation and afterward separated the entire sound document into parts I.e., windows and made profile for each time stamps and afterward these pieces of sound record are passed to show for additional handling

VII. ALGORITHM

The picked calculation is a feed-forward brain network utilizing an old style inclination plummet calculation with a negative log-probability as cost work. The brain network engineering is the accompanying. There are twelve information credits, which connect with the twelve semi-tones of the PCP vector addressing the harmony. The brain network yields a vector of 10 qualities, comparing to the result neurons, every one being the likelihood of the identified harmony to be given from the relating harmony. The last settings of the brain network were upgraded by a 10-crease cross-approval on the learning information base.

1. Analysis of Audio File –

In Audio Analysis after the quiet is eliminated, the sound sign is imagined based on the recurrence range and the positive and negative frequencies of the sound sign. Different boundaries like recurrence inspecting, the length of sign, channels and the time venture between tests is dissected. Figure underneath shows the sound record examination.

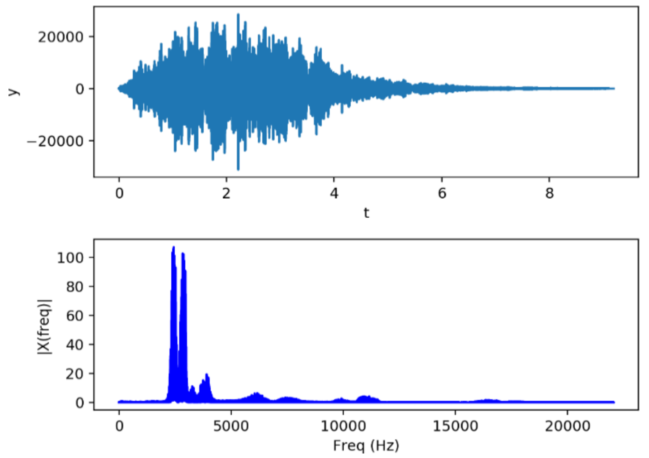

Figure 3: Audio file analysis for frame rate and frequency detection

2.Analysis of audio file –

Determine the sampling rate 'fsrate' based on signal 's' value from the silence removed input file I'. Input: The silence removed audio file I'.
Fsrate, s ←scipy.io. wavfile. read (I')
Secs←len (s. shape)
ts←1.0/fsrate
FFT←scipy.fft (secs, ts)
freq←scipy. fftpack. fftfreq (secs, ts)

3. Pitch Profile Classifier for Identification

Determine the pitch of the profile based on the Input audio file and the I'.

Input: Analyzed Audio Data 'A'
Fsrate, s ←scipy.io.wavfile.read(I')
pcp←GeneratePcp (fsrate, s)
PcpNormalised←NormalizePcp (pcp)

4. GeneratePcp () and NormalizePcp (pcp)-

Determine the pitch based on sampling rate and frequency
GeneratePcp () –
Input: sampled frequency
for sample =1 to N:
if (freq(sample) == 0):
return -1
else:
return (12*log2 (fs*1) / (N * freq(sample))) % 12

NormalizePcp (pcp) –
Forpcp [index=0 to N samples]:
PcpNormalized←pcp[index]/sum(pcp)
return PcpNormalized

VIII. CONCLUSION

As indicated by the market overview, we observed that there are two-three applications in market for the proposed issue articulation likewise, as per writing study, there were individuals who attempted to find the answers for harmony acknowledgment issue proclamation yet this large number of arrangements and applications slacked in giving clients, straightforward highlights for novices in area of music, blunder less harmony sheet as certain destinations and applications are dealt with physically, and so forth. Henceforth all the connected work, project scope in report assisted with clearing the vision for project. The proposed task will assist with eliminating the disadvantages we found in review by giving elements to clients like simple to-peruse harmony sheet production of harmony sheet for your tunes, playing and auto looking over harmony sheet as per clients' speed, making playlist as indicated by clients decision. These elements are applied in the finished result that is web application utilizing AI in the backend thus they will accompany least mistakes.

We will probably utilize methods to recognize a word reference of more modest parts that can give the most proficient inclusion of huge music data sets. While applying every one of the elements in project we attempted to deal with non-useful prerequisites for the undertaking like unwavering quality, convenientce, execution, adaptability, responsiveness, and so forth as the venture will be utilized on various conditions, with shifting responsibility subsequently task ought to have the option to give clients fast reaction, dependable highlights and great execution in various circumstances. As the venture of harmony acknowledgment of obscure melodies requests the progressions in the improvement stage as indicated by necessities of clients consequently from all the examination, we presumed that the stream project work will follow iterative model because of which toward the finish of task we will actually want to give the essential elements and undertaking targets to clients on time and with precision. Taking into account every one of the client's necessities, programming prerequisites we attempted to foster AI model utilizing brain network utilizing dataset which involves harmonies present in boisterous, commotion free melodies which gave productive forecast to obscure tunes and consequently to the issue articulation.

REFERENCES

[1] J.-J. Aucouturier and F. Pachet. Music similarity measures: What's the use? In Proceedings of the 3rd International Conference on Music Information Retrieval (ISMIR), pages 157–163, 2002

[2] T. Bertin-Mahieux, D. Ellis, B. Whitman, and P. Lamere. The million song dataset. In Proceedings of the 12th International Conference on Music Information Retrieval (ISMIR), 2011.This paper presents a method based on machine learning techniques, using a feed-forward neural network, for

[3] M. Carre. Systèmes de Recherche de Documents Musicaux par Chantonnement. PhD thesis, École Nationale Supérieure des Télécommunications, Paris,2002.

[4] M. Deza and E. Deza. Encyclopedia of Distances.Springer, 2009.

[5] A. Egorov and G. Linetsky. Cover song identification with if-f0 pitch class profiles. In MIREX extended Abstract, 2008.

[6] D. Ellis and C. Cotton. The 2007 labrosa cover song detection system. In MIREX Extended Abstract, 2007.

[7] D. Ellis and G. Poliner. Identifying cover songs with chroma features and dynamic programming beat tracking. In Proceedings of the IEEE Int. Conf. On Acoustics, Speech, and Signal Processing (ICASSP), 2007.

[8] A. Flexer, D. Schnitzer, M. Gasser, and G. Widmer.Playlist generation using start and end songs. In Proceedings of the International Conference on Music Information Retrieval (ISMIR), pages 173–178, 2008.

[9] T. Fujishima. Realtime chord recognition of musical sound: a system using common lisp music. In Proceedings ot the International Computer Music Conference (ICMC), pages 464–467, 1999

[10] E. Gomez and P. Herrera. Automatic extraction of tonal metadata from polyphonic audio recordings. In Proceedings of the International Conference: Metadata for Audio, 2004.

[11] E. Gomez and P. Herrera. The song remains the same; identifying versions of the same song using tonal descriptors. Proceedings of the International Conference on Music Information Retrieval (ISMIR), pages 180–185, 2006.

[12] C. Harte and M. Sandler. Automatic chord identifcation using a quantised chromagram. In Proceedings of the 118th Audio Engineering Society Convention(AES), 2005.

[13] T. Hastie, R. Tibshirani, and J. Friedman. The elements of statistical learning: data mining, inference, and prediction. Springer Series in Statistics.Springer, second edition, September 2009

[14] J. Jensen, M. Christensen, D. Ellis, and S. Jensen. A tempo-insensitive distance measure for cover song identification based on chroma features. In Proceedings of the IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP), pages 2209–2212,2008.

[15] S. Kim and S. Narayanan. Dynamic chroma feature vectors with applications to cover song identification. In Proceedings of the IEEE Workshop on Multimedia Signal Processing (MMSP), pages 984–987, October 2008.

[16] K. Lee. Automatic chord recognition using enhanced pitch class profile. In Proceedings of the International Computer Music Conference (ICMC),pages 306–313, 2006

[17] M. Müller, F. Kurth, and M. Clausen. Audio matching via chroma-based statistical features. In Proceedings of the International Conference on Music Information Retrieval (ISMIR), pages 288–295, 2005.

[18] J. Serra, E. Gómez, and P. Herrera. Audio cover song identification and similarity: Background, approaches, evaluation, and beyond. In Advances in Music Information Retrieval, pages 307–332. Springer, 2010.

[19] A. Sheh and D. Ellis. Chord segmentation and recognition using EM-trained hidden markov models. In Proceedings of the International Conference on Music Information Retrieval (ISMIR), pages 121–124, 2003.

[20] L. Xiao, L. Liu, F. Seide, and J. Zhou. Learning a music similarity measure on automatic annotations with application to playlist generation. In Proceedings of the International Conference on Acoustics, Speech and Signal Processing, pages 1885–1888,2009.