International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 09 | Sep 2020 www.irjet.net p-ISSN: 2395-0072

MUSIC RECOMMENDATION SYSTEM FROM PLAYLIST BASED ON ECG SIGNAL USING NN CLASSIFIER

D J Hani Mary Sheniha[1], Anitha Loraine A[2], Madhumathi S[3]

[1]Assistant professor, Department of Computer Science and Engineering, JEPPIAAR SRR Engineering College, Tamil Nadu, India
[2]Department of Computer Science and Engineering, JEPPIAAR SRR Engineering College, Tamil Nadu, India
[3] Department of Computer Science and Engineering, JEPPIAAR SRR Engineering College, Tamil Nadu, India

---

Abstract - *The project presents emotion recognition from ECG signals based on feature analysis and NN-classifier. Automatic emotion recognition plays an important role in HCI systems. After getting the emotion results, we can play the song according to their moods. This emotion recognition system will help the people who are alone and distress in hospitals.*

***Key Words*: Emotion recognition, Pre-processing, Signal Compression, Feature Extraction, NN Classifier.**

1. INTRODUCTION

Listening to music can changes the person's mood and they feel good. An automatic music recommendation system can be used to select a music that is appropriate for a patient's changing mood.

Emotion recognition is usually based on four crucial steps: (1). original affective data acquisition; (2). affective feature extraction; (3). affective feature subset selection; (4). emotion classification. However, whether the target emotion is elicited is the key issue to gain corresponding affective data. Then it can influence the following work. Thus affective data acquisition is the foundation of other works. One of the interesting topics in Human Computer Interaction (HCI) is to understand human emotions and classify emotions and this is one of the issue towards emotional intelligence. Research on human emotions has been carried out with interest in various fields, and emotion recognition methods of a driver in automotive research have been presented, recently. Emotion recognition using biomedical signals can acquire a signal easily from a sensor, and this biomedical signal has a natural emotional state because it is under the control of the autonomic nervous system. Therefore, it has an advantage that it can be used for emotional recognition. Affective variations can cause physical changes. Some emotion features may hide in the ECG signals. A multichannel physical signal recording machine, namely, BIOPAC System MP150, was applied to record original ECG signal from body surface. During the recording time, subjects were allowed to watch film clips in non-interference environment to arouse certain emotion without much discomfort. The sampling frequency was set to 200Hz.

1.1 DIGITAL SIGNAL PROCESSING

Digital Signal Processing (DSP) is concerned with the representation of discrete time signals by sequence of numbers or symbols and the processing of these signals. Digital signal processing and analog signal processing are subfields of signal processing.DSP includes subfields like: audio and speech signal processing, sonar and radar signal processing, sensor array processing, spectral estimation, statistical signal processing, digital image processing, signal processing for communications, control of systems, biomedical signal processing, seismic data processing.

DSP manipulates different types of signals with the intension of filtering, measuring or compressing and producing analog signals. Analog signals differ by taking information and translating it into electric pulses of varying amplitude, whereas digital signal information is translated into binary format where each bit of data is represented by two distinguishable amplitudes. Another noticeable difference is that analog signals can be represented as sine waves and digital signals are represented as square waves. DSP can be found in almost any field, whether it's sound reproduction, radar and sonar, medical image processing or telecommunications essentially any applications in which signals are being compressed and reproduced.

The digital signal process takes signals like audio, voice, video, temperature or pressures that have already been digitized and then manipulates them mathematically. This information can then be represented discrete time, discrete frequency and other discrete forms so that the information can be digitally processed. An analog-to-digital converter is needed in the real world to take analog signals and convert them into 0's and 1's for digital format.

The main task of ECG signal preprocess and filter signals prior to analysis and then explore and extract features for data analytics and machine learning applications. Then Analyze trends and discover patterns in signals. Finally visualize and measure time and frequency characteristics of signals.

The goal of DSP is usually to measure, filter and/or compress continuous real-world analog signals. The digital signal processing involves a large number of operations being performed quickly on a set of data over and over again. These operations are carried out by the DSP algorithms. Signals are constantly converted from analog to digital, processed digitally, and then converted again to analog. Consider an example of a voice recording where voice is an analog signal which is recorded through a microphone. The first step is to convert this voice signal to digital form using an ADC (analog-to-digital converter). This converted data, now contains a stream of numbers which is manipulated digitally by the DSPs and converted back to an analog signal with the help of a Digital-to-analog converter. This analog signal is again a sound that could be played back.

1.2 NEURAL NETWORK CLASSIFIER

Neural Network (NN) and General Regression Neural Networks (GRNN) have similar architectures, but there is a fundamental difference: networks perform classification where the target variable is categorical, whereas general regression neural networks perform regression where the target variable is continuous. If select a NN/GRNN network, DTREG will automatically select the correct type of network based on the type of target variable.

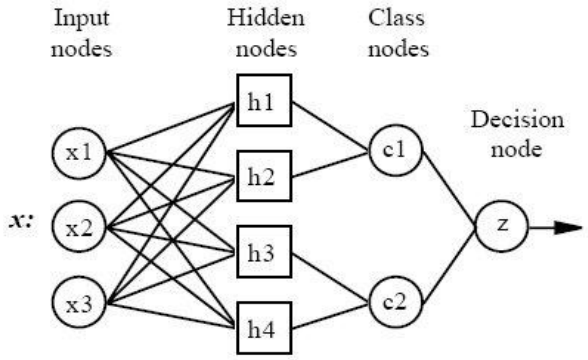

Fig -1: Neural Network Classifier

Neural Network contains has four layers such as Input layer, Hidden layer, Pattern layer and Decision layer. In Input layer, there are one neuron in the input layer for each predictor variable. In the case of categorical variables, *N*-1 neurons are used where *N* is the number of categories. The input neurons (or processing before the input layer) standardize the range of the values by subtracting the median and dividing by the interquartile range. The input neurons then feed the values to each of the neurons in the hidden layer.

In Hidden layer, one neuron for each case in the training data set. The neuron stores the values of the predictor variables for the case along with the target value. When presented with the *x* vector of input values from the input layer, a hidden neuron computes the Euclidean distance of the test case from the neuron's center point and then applies the RBF kernel function using the sigma value(s). The resulting value is passed to the neurons in the pattern layer.

In Pattern layer, the next layer in the network is different for NN networks and for GRNN (General Regression Neural Network) networks. For NN networks there is one pattern neuron for each category of the target variable. The actual target category of each training case is stored with each hidden neuron; the weighted value coming out of a hidden neuron is fed only to the pattern neuron that corresponds to the hidden neuron's category. The pattern neurons add the values for the class they represent (hence, it is a weighted vote for that category).For GRNN networks, there are only two neurons in the pattern layer. One neuron is the denominator summation unit the other is the numerator summation unit. The denominator summation unit adds up the weight values coming from each of the hidden neurons. The numerator summation unit adds up the weight values multiplied by the actual target value for each hidden neuron.

The Decision layer is different for NN and GRNN networks. For NN networks, the decision layer compares the weighted votes for each target category accumulated in the pattern layer and uses the largest vote to predict the target category. For GRNN networks, the decision layer divides the value accumulated in the numerator summation unit by the value in the denominator summation unit and uses the result as the predicted target value. Neural Network contains some of the advantages:

- Neural networks are broadly used applications for financial operations, enterprise planning, trading, business analytics and product maintenance.
- Neural networks have also gained widespread adoption in business applications such as forecasting and marketing research solutions, fraud detection and risk assessment.
- It is usually much faster to train a NN/GRNN network than a multilayer perceptron network.
- NN/GRNN networks often are more accurate than multilayer perceptron networks.
- NN/GRNN networks are relatively insensitive to outliers (wild points).
- NN networks generate accurate predicted target probability scores.
- NN networks approach Bayes optimal classification.

2. EXISTING SYSTEM

In machine learning, Support Vector Machines(SVM) are supervised learning models with associated learning algorithms that analyse data used for classification and regression analysis and also Principle Component Analysis (PCA) is used in machine learning for predictive models and data analysis. A geometric method is a branch of mathematic concerns with shape, size and properties of space. Principle Component Analysis (PCA) is a statistical procedure that uses an orthogonal transformation to convert a set of observation of possible correlated variable into a set of value of linearly uncorrelated value.

2.1 RELATED WORKS

Yui Matsumoto, RyosukeHarakawaand Miki Haseyama has planned a way for music video recommendation is given. The contribution of square measure two-fold.The planned technique constructs a network, that not solely represents relationships between music videos and users however additionally captures multi-model option of music videos. This Music videos permits a collaborative (cooperative) use of multi-model options like audio, visual, textual options and multiple social information that may represent a relationships between music videos and video hosting services. A completely unique theme for the local and global structures considers the various link prediction of the network (LP-LGSN) and it is newly derived by fusing multiple link prediction scores based on both local and global structures. By victimization the LP-LGSN is used to predict the degrees to that users need music videos, the planned technique will suggest users' desired music videos. The final results for a real-world dataset made from YouTube-8M show the effectiveness of the planned technique.

Kunhui Lin, ZhentuanXu and Jie Liu has projected a way for mobile web and streaming media technology, digital music has been accepted by many of us. However, within the face of the huge music information on the web, users wish to search out favorite music is like searching for a needle in an exceedingly stack via ancient music classification techniques. An improved algorithmic rule for customized music recommendation supported tag data is projected. Firstly, to use the improved user-based collaborative (cooperative) filtering algorithmic rule to contend with the user's long preferences. Secondly, in accordance with the user-tag-music relationships, obtaining the music that related to the user via recommendation algorithmic rule supported bipartite graph. Experiments on real dataset demonstrate the potency and effectiveness of the improved algorithmic rule, recoupling results and have sensible significance.

Markus Schedl has planned strategy for music genre were found to be underrepresented on social media and genre streaming platforms, they represent a crucial target cluster for music recommender systems. Thus specialise in this cluster of listeners and investigate a good vary of advice approaches and variants for the task of music creator recommendation. At intervals the cluster of music genre listeners, additional assess categorizing users in line with demographics and temporal music consumption behaviour. To report the results of preliminary recommendation experiments and insights gained for the attender cluster into consideration.

S.Kuchibhotla, H.D.Vankayalapati, and K. R. Anne has projected a technique for Feature Fusion plays an vital role in speech feeling recognition to boost the classification accuracy by combining the foremost widespread acoustic options for speech feeling recognition like energy, pitch and Mel frequency cepstral coefficients. But the performance of the system isn't best as a result of the procedure quality of the system, that happens thanks to high dimensional related to feature set when feature fusion. A two stage feature choice methodology is projected. In first stage feature choice, applicable options area unit elect and consolidated along for speech feeling recognition. In second stage feature choice, best feature set choice techniques [sequential forward choice and consecutive floating forward selection (SFFS)] area unit want to eliminate the curse of spatiality drawback thanks to high dimensional feature vector when feature fusion. Finally the emotions area unit classified by mistreatment many classifiers like Linear Discriminant Analysis (LDA), Regular Discriminant Analysis (RDA), Support Vector Machine (SVM) and K Nearest Neighbor (KNN).

Anbang Yao and Shan Yu has projected how for representing a face appearance. Along with context deal with the face image as a group of small facial regions, a different new face representation approach coined spatial option interdependence matrix (SFIM). Not like other classical face descriptors that generally use a hierarchically organized structured or a sequentially concatenated structure to verify the spatial layout choices extracted from local regions, SFIM is attributed to the exploitation of the underlying feature interdependences regarding native region among a class specific face. According to SFIM, the face image is projected onto an undirected connected graph and it is associated extremely manner that explicitly encodes option interdependence-based relationships between native regions. To calculate the pair-wise relation strength as the result of the weighted discrepancy between two feature sets extracted in a associate extremely hybrid feature space fusing histograms of intensity, local binary pattern and orientated gradients. To realize the goal of face recognition, our SFIM-based face descriptor is embedded in three all totally different recognition frameworks, namely nearest neighbor search, subspace-based

classification, and linear optimization-based classification. Full experimental results on four well-known face databases and comprehensive comparisons with the state-of-the-art results unit of measurement provided to demonstrate the efficacy of the proposed SFIM-based descriptor.

3. PROPOSED WORK

3.1 OVERVIEW

The Music Recommendation System is the most probable way for the people who are suffocated silence and loneliness in ICU as well as in the hospitals which cheer them up during the period of loneliness they go through in the hospitals. This system helps them to be out of the anxiety that they face .The ECG signal from the person is taken as input to the pre-processor, all the noise along the original wave have been detected using the discrete wavelet transform, the compressed wave from the DWT is filtered using low pass and high pass filters and then they are analyzed using the NN classifier which classifies the signal and play the music according to the person's mood, this would help them to overcome dropdown in their heart beat rate.

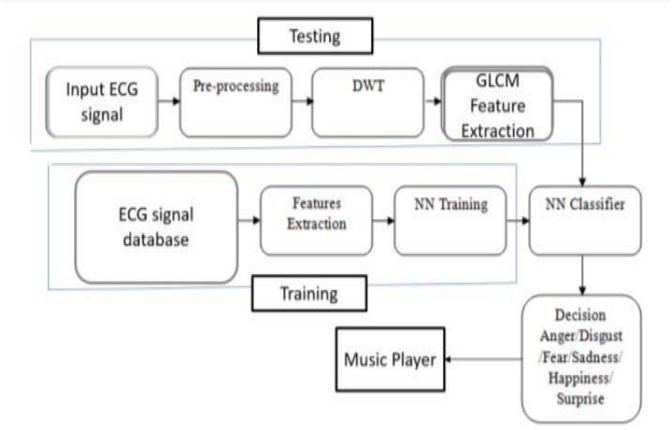

Fig -2: Music Recommendation System using ECG signal

3.2 MODULAR DESIGN

The recognition system involves emotion detection, feature extraction, selection and finally classification. These features are useful to distinguish maximum number of samples accurately. Performing at low time complexity and accuracy will increase the operational function.

3.2.1 Input Signal Pre-processing

Pre-processing eliminates the noise and produces a cleaner, more easily interpretable ECG. Attempts to remove, rather than model and data variability. These may include cleaning, integration, transformation and reducing process of improvement. The pre-processing techniques are contrast stretching, global thresholding, histogram equalization, log transformation and power law transformation. The important function of pre-processing are noise removal, reduction of blurring, increasing contrast and providing more detail. Unwanted signal is removed.

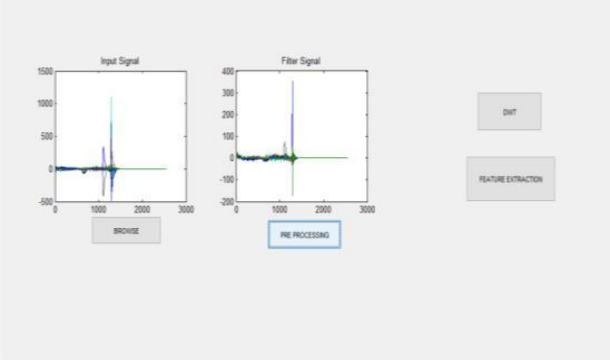

Fig -3: Input Signal Pre-processing

3.2.2 ECG Signal Compression

Discrete Wavelet Transform decomposes a signal into a set of basis function called wavelet. DWT compresses the original signal. The compressed signal ends up with the same number of coefficients as the original signal. DWT contains high-quality signal approximation. The Discrete Wavelet Transform are equivalent to discrete filter banks. The signal is first filtered by a low pass and a high pass filter to yield low pass and high pass sub bands. In the DWT, the filter outputs are down sampled at each successive stage.

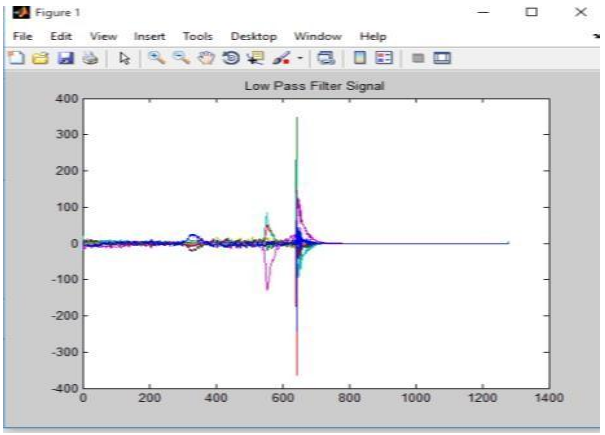

Fig -4 : Low Pass Filter Signal

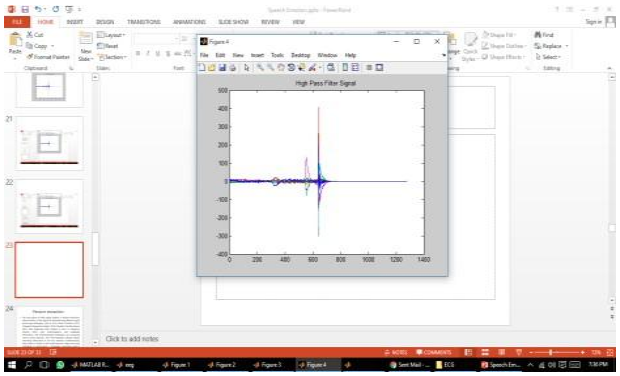

Fig -5 : High Pass Filter Signal

3.2.3 GLCM Feature Extraction

GLCM extracts the textural features. The Gray Level Co-occurrence Matrix is a method which extracts the features like mean, entropy, skewness and variance. To create a GLCM, the GLCM function is employed. The grey level co-occurrence matrix function denotes a grey-level co-occurrence matrix (GLCM) is used to calculate a pixel with the intensity (gray-level) value i denotes during a particular spatial relationship to a pixel with the worth. Usually, the spatial relationship is defined because the pixel of interest and thus the pixel to its immediate right (horizontally adjacent), but you'll specify other spatial relationships between the two pixels. Each element (i,j) within the resultant GLCM is just the sum of the amount of times that the pixel with value i occurred within the specified spatial relationship to a pixel with value j in the input signal.

Mean: For a variable vector *A* creates of *N* scalar observations, the mean is defined as

$$\mu = \frac{1}{N}\sum_{i=1}^{N} A_i,$$

Homogeneity: Measures the closeness of the distribution of various elements in the GLCM to GLCM diagonal.

Homogeneity = sum (sum (p(x, y)/ (1 + [x-y])))

Where, x and y – Coordinates, p(x, y) is the count of the each co-occurrence pair

CORRELATION: Measuring the linear dependency of the grey levels of the neighbouring pixels. Therefore it measures the grey tone linear dependence.

ENERGY: Energy refers to some value or vector of values that are used to minimize, maximize or optimize.

Category	Signal 1	Signal 2
Mean	0.1092	0.1738
Skewness	0.7660	0.7453
Kurtosis	0.0053	0.0042
Correlation	10.1127	6.0717
Energy	0.8590	0.5604

Table - 1: Feature Extraction

3.2.4 Emotion Classification

Neural networks are predictive models loosely supported on the action of biological neurons. Neural network is that the best tool in recognition and discrimination between different sets of signals. The network has an input layer (on the left) with three neurons, one hidden layer (in the middle) with three neurons and an output layer (on the right) with three neurons.

Neural Network classifier is used as a tool for emotion classification. NN classifier helps to cluster the data according to the similarities among the example inputs and the dataset. Based on similarities is found, emotion is classified.

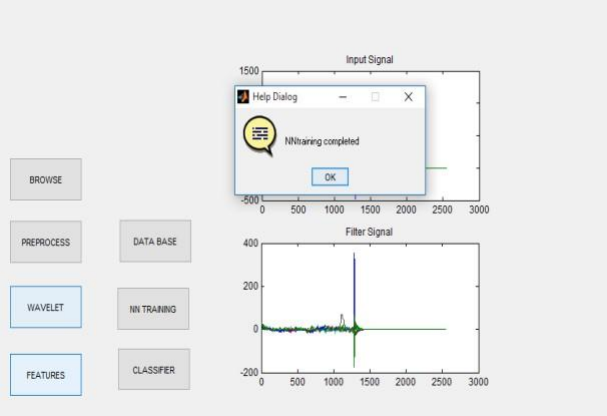

Fig -6 : Emotion Classification

3.2.5 Emotion Based Music Play

Based on the emotion classified from NN classifier the music is played. The music is classified into various genres which highlighted into country, jazz, hip-hop, disco, pop. According to the wave signal extracted from the emotion, different music is played.

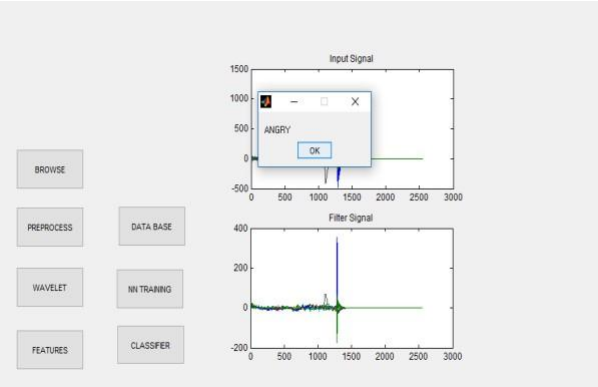

Fig -7 : Emotion Based Music Play

4. CONCLUSION

Music Recommendation System will bring the bright side of the lonely person's in the hospital who have no one to take care of them. Hearing an positive vibe will bring an extra belief in the life of the patients. This emotion identifier is surprisingly effective life lighter for lonely ones

REFERENCES

[1]Yui Matsumoto, RyosukeHarakawa and Miki Haseyama, "Music Video Recommendation Based On Link Prediction Considering Local and Global Structures of a network", IEEE Transactions on Network Analysis and Social meta-data, Vol. 7, no. 4, pp. 14-17, July 24,2019.

[2 KunhuiLin ZhentuanXu and Jie Liu, "Personalized music recommendation algorithm based on tag information", Computer Engineering and Applications, Vol. 48, no.18, pp. 119-122, July 2016.

[3]Markus Schedl, "Towards Personalizing Classical Music Recommendations", Computer Vision and Image Understanding, Vol.28, no.14, pp.26-37, Nov 2015.

[4]S. kuchibhotla, H.D.Vankayalapati and K. R. Anne, "An optimal two stage feature selection for speech emotion recognition using acoustic features",International Journal of Speech Technology, Vol. 19,no. 4, pp. 1-11,May 2014.

[5]Anbang Yao and Shan Yu, "Robust Face Representation Using Hybrid Spatial Feature Interdependence Matrix", Pattern Analysis and Machine Intelligence, Vol. 22, no. 88, pp. 33-58, August 2013.