International Journal of Advanced Trends in Computer Applications
www.ijatca.com

Emotion Recognition Using SVM and NN

[1]Shivani Goel, [2]Dr.Kiran Jyoti
[1]M Tech student, Information Technology
Guru Nanak Dev Engineering College, Ludhiana
[2]Deptt.of Information Technology
Guru Nanak Dev Engineering College, Ludhiana
[1]goel109@gmail.com, [2]kiranjyoti@gmail.com

Abstract: *Feelings assume a greatly imperative part in human mental life. It is a medium of articulation of one's point of view or his mental state to others. It is a channel of human mental portrayal of one's emotions. Feelings are a key grammatical form. Naturally recognizing feeling in a recording can improve human PC communication. It likewise empowers different sorts of investigations, for example, scan for paralinguistic phenomena, the genuineness of the speaker, and so on. In this proposition, we are going to utilize neural system, SVM and HMM strategy on discourse that is concentrate by highlight extraction technique. The characterization execution is taking into account separated highlights utilizing different classifiers. There may be sort of feelings amid discourse like happy, sad, aggressive and fear..In this work these four feelings are going to recognize. On premise of these feelings at long last we attain to on a conclusion with exactness results. The entire recreation has been occurred in MATLAB environment.*

Keywords: *BPNN, Classification, Emotion Recognition, Speech, SVM, HMM.*

I. Introduction

Discourse feeling acknowledgment means to consequently recognize the current passionate condition of a person from his or her voice. It is in light of top to bottom examination of the era component of discourse sign, separating some trademark parameters which contain passionate data from the speaker's voice, utilizing these parameters and taking proper example acknowledgment techniques to recognize enthusiastic states from discourse.

Presently, Automatic Discourse Emotion Recognition is an extremely dynamic examination theme in the Human Computer Interaction (HCI) field and has an extensive variety of uses. For separation learning, recognizing understudies' feeling convenient and making suitable treatment can improve the nature of instructing. In programmed remote call focus, it is utilized to convenient recognize clients' disappointment. It is likewise used to help clinical analysis or to play feature diversions. The examination of programmed discourse feeling acknowledgment, not just can advance the further improvement of PC innovation, it will likewise incredibly upgrade the productivity of individuals' work and study, and help individuals tackle their issues all the more proficiently. It will likewise further enhance our lives and enhance the personal satisfaction.

Lately, a lot of exploration has been done to perceive human feeling utilizing discourse data. Numerous discourse databases were manufactured for discourse feeling exploration, for example, BDES (Berlin Database of Emotional Speech) that is German Corpus and built by Department of acoustic innovation of Berlin Technical University [1], DES (Danish Emotional Speech) that is Danish Corpus and secured by Aalborg University, Denmark [2], SES (Spanish Emotional Speech) that is Spanish Corpus. There are likewise some Mandarin Affective Speech Databases, for example, MASC (Mandarin Affective Speech) and MESC (A Mandarin Emotional Speech Corpora) that is recorded by Tsinghua University, Taiwan. Numerous analysts have proposed essential discourse highlights which contain feeling data, for example, vitality, pitch recurrence [2], formant recurrence [3], Linear Prediction Coefficients (LPC), Linear Prediction Cepstrum Coefficients (LPCC), Mel-Frequency Cepstrum Coefficients (MFCC) and its first subordinate [4].

Moreover, numerous specialists investigated a few order techniques, for example, Neural Networks (NN) [5],

Gaussian Mixture Model (GMM), Hidden Markov model (HMM) [6], Maximum Likelihood Bayesian classifier (MLC), Kernel Regression and K-closest Neighbors (KNN) and Support vector machines (SVM) [7].

II. CLASSIFICATION OF EMOTIONS

How to classify emotions is an interesting and difficult issue. Researchers on emotion recognition differ on number of categories to use. Some classification systems which have been used include:

- neutrality, happiness, boredom, sadness, anger, fear [2]
- anger, fear, Badness, joy, disgust [3]
- neutrality, happiness, sadness, anger, fear, boredom, disgust [4]
- fear, anger, sadness, happiness [5]

After examining these examples, we have selected to use following four emotional states in this study:
Sadness, happy, anger and aggressive.

III. SYSTEM IMPLEMENTATION

In this proposed work we have utilized discourse tests for the database. In the database we discover properties of the discourse signs and after that we store them into the database. The inquiry comes that how we are going to store many records in the database [8]. The strategy would be as per the following. As a matter of first importance we would bring the properties of the voice tests. Every one of those properties which are obliged would be figured and after that it would be put away into a cluster. The cluster would proceed onward as the documents would move. We would bring the highlights and would take the normal by the end and after that store them into the database for every classification of the voice which we have taken i.e. HAPPY, SAD, ANGRY AND AGGRESSIVE.

A. Voice Files

The voice documents are the records which would be prepared for the highlight extraction.

B. Properties

When we would prepare the voice documents their properties would be brought. For the highlight extraction there are a few calculations which can be utilized [9]. In our methodology we have utilized HMM calculation for the preparation reason.

C. Training

In the preparation segment we would be taking different voice tests of every single classification taken for the order. In this situation, we would be getting properties of every voice test and in the wake of placing them into a cluster. We would be putting away the normal of every property of every area into the database. To accomplish this specific undertaking, we would be utilizing HMM calculation. The preparation segment guarantees that the database gets prepared appropriately so that at the season of testing it creates broad results [10]. The highlights of preparing are as per the following

a. Maximum Frequency: The most extreme recurrence of a document is the worth which we get at the crest on a recurrence map. At the point when so ever we put a voice test over the time and recurrence design, the greatest crest is known as the most extreme recurrence of the voice test.

b. Minimum Frequency: The base recurrence of a record is the worth which we get at the top on a recurrence map. At the point when so ever we put a voice test over the time and recurrence design, the base top is known as the base recurrence of the voice test.

c. Average Frequency: The normal recurrence can be computed utilizing two procedures. The primary method is to include all the recurrence tests and after that gap the whole entirety with the aggregate number of recurrence. The second strategy is an exceptionally moral system in which we can include the base recurrence and the most extreme recurrence and after that we can isolate them by two.

Avg Frequency = (Minimum recurrence + Maximum Frequency)

d. Spectral Roll off: The unearthly move off as far as advancement can be said as the contrast between the greatest recurrence contrasts with the adjoining recurrence. The position of the recurrence (max) can be put away into an exhibit and comparable of the nearby hub and afterward the distinction can be computed.

e. Noise Level: Ethically the commotion level is the additional number of bits which has been included into the voice test. On the off chance that the commotion is uniform then the clamor level can be computed by taking the distinction of every recurrence test and the edge of the voice test. There are two classifications of commotion level: uniform clamor and non-uniform commotion.

f. **Uniform Noise:** Uniform clamor is the commotion which is all the while same everywhere throughout the voice test.

g. Non Uniform Noise: The non-uniform clamor does not stay consistent everywhere throughout the example.

h. Pitch: It is the normal estimation of the whole voice test.

i. Spectral Frequency: The ghostly recurrence is the recurrence of the voice pitch beside the most astounding voice test.

D. Testing

At the season of testing we would be utilizing a combinational calculation utilizing the SVM and NEURAL nourish forward strategy. In this part, we would be binarizing the spared information of the database and would give it to the NEURAL classifier. On the premise of the spared database, the neural classifier would coordinate the properties with the transferred record and would deliver an outcome. The characterization would be done on the premise of 4 classes: content, dismal, furious and forceful [11].

F. Feature Extraction Algorithm

SVM (Support Vector Machine) Algorithm:

1. Firstly train the designed system using SVM

```
SvmStruct = svmtrain(pbest(Y,:),group(Y,:));
```

2. Classify it using SVM classify command

```
Classes = svmclassify(svmStruct,pbest(X,:));
```

3. Evaluate the class performance

```
Classperf(cp, classes, Y);
```

Neural Network Algorithm:

1. Create network

```
Net = network;
```

2. Set number of inputs

```
net.numInputs = 2;
```

3. Assign 2 to input size

```
net.inputs{1}.size = 1;
net.inputs{2}.size = 1;
```

4. Add 1 layer to network

```
net.numLayers = 1;
```

5. Assign number of neurons in layer

```
net.layers{1}.size = 1
```

6. Connect input to layer 1

```
net.inputConnect(1) = 1;
net.inputConnect(2) = 1;
net.biasConnect(1) = 1;
```

7. Set network init function

```
net.initFcn = 'initlay';
```

8. Set network training function

```
net.trainFcn = 'trainc';
```

9. Set network performance evaluation

```
net.performFcn = 'mae';
```

10. Evaluate the network

```
View(net) ;
```

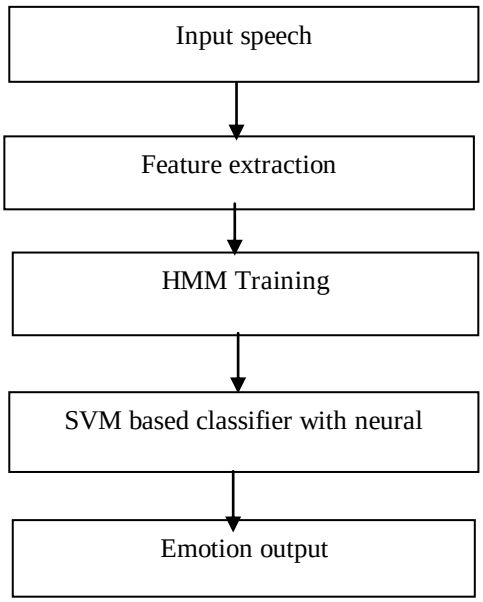
```
Net = train(net,[0 0 1 1;0 1 0 1],[0 0 0 1])
```

A. Validation

Neural Networks is one of the most advanced classifiers in the testing category. The neural has a feed forward method. The feed forward method takes one input as a training sample and another input as the target sample. The input sample is the data stored in the database on the basis of all the features which have been extracted at the time of training. If P is the input sample then P would be defined as P =sum (all. Features (input)); In the same manner, there would be the testing feature or the target sample. The target sample would be the scenario which would be fixed and would be provided as an input [12].

IV. SIMULATION MODEL

The whole simulation has been taken place in MATLAB environment with the flow of following sections:



V. IMPLEMENTATION AND RESULTS

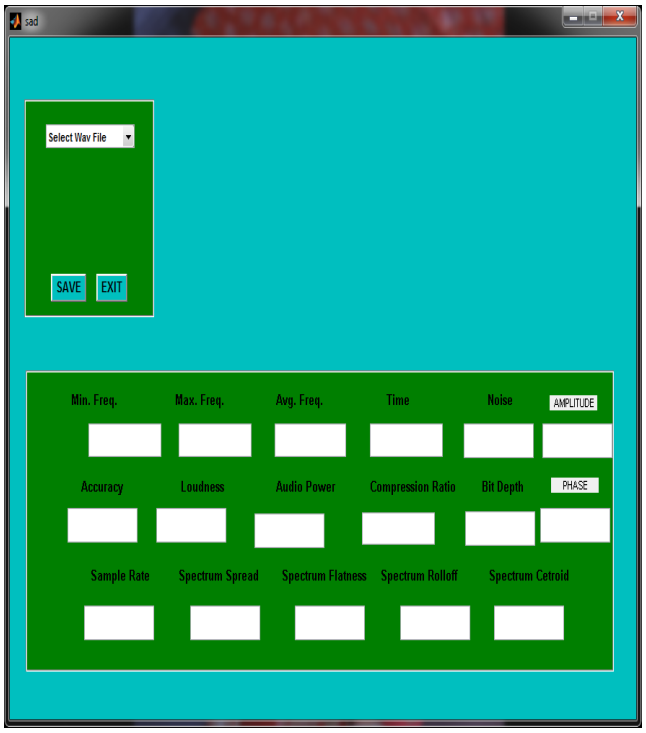


Figure 1: Graphical User Interface panel

Figure 1 shows the GUI panel for the proposed system. The above figure shows the features in the panel like maximum frequency, average frequency etc. which are extracted after feature extraction process.

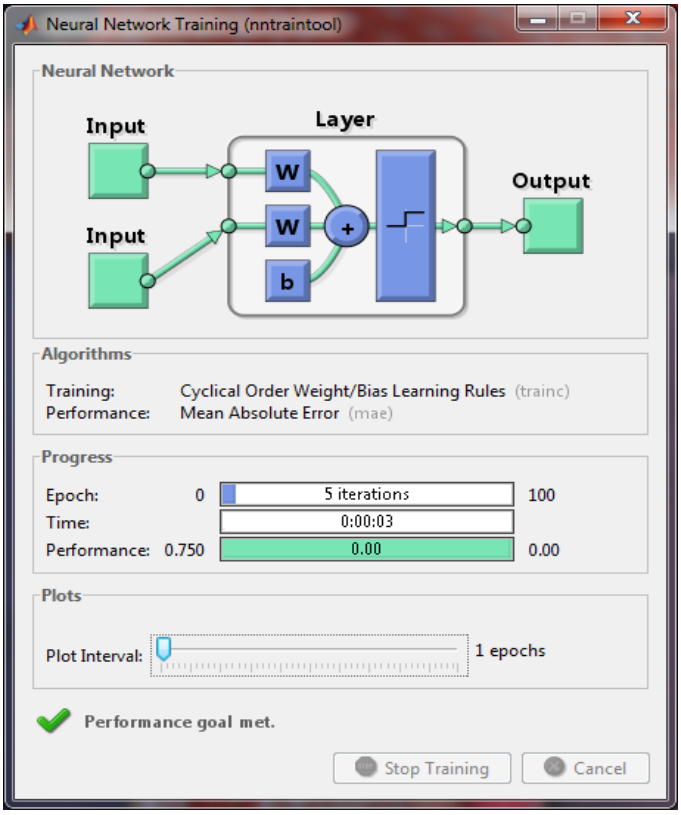


Figure 2: Neural Network

The Figure 2 shows the neural network having layer structure which is used to train the network. The network is to train the system having number of epochs called iterations. The network shows the input layer, hidden layer and output layer.

In this validation part, the results are compared in terms of accuracy between SVM and Neural Network.

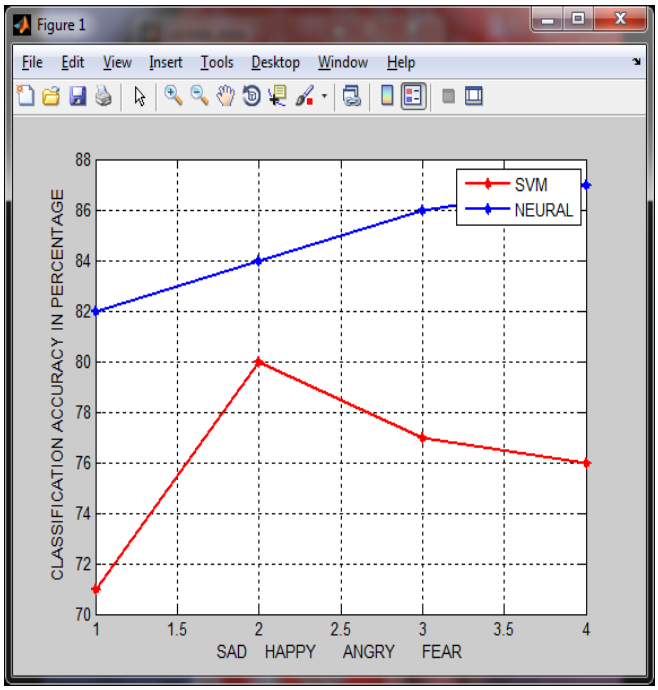


Figure 3: Validation result from SVM and NN

The above figure shows that the neural is more efficient than Support Vector Machine having more accuracy than SVM.

VI. CONCLUSION AND FUTURE SCOPE

Undergoing the estimated procedure, we expect our conclusion to be a better accurate system for the analysis of the audio files to detect the emotions in the field of clustering. We expect the accuracy to be increased by 2 to 5 percent in comparison with the ANN. SVM combined with HMM and neural is expected to work in better manner because the training set created with the help of SVM and HMM puts a strong emphasis in searching into the inner clusters of the files.

Although the results are efficient enough to provide solution, but our work provides a lot for the future researchers. The current scenario is not suitable for noisy audio files, if we increase the noise level of audio files the current scheme might fail to produce efficient results.

In future, work can also be done to create more groups into the inner cluster of the files stored so that the searching becomes easy. The future parameters can add the time slots of the frequencies at which the frequencies are consistent.

REFERENCES

[1] http://www.expressive-speech.net/. Berlin emotional speech database

[2] D. Ververidis, C. Kotropoulos, and I. Pitas, "Automatic emotional speech classification", in Proc. 2004 IEEE Int. Conf. Acoustics,Speech and Signal Processing, vol. 1, pp. 593-596, Montreal, May2004.

[3] Xiao, Z., E. Dellandrea, Dou W.,Chen L., "Features extraction and selection for emotional speech classification". 2005 IEEE Conference on Advanced Video and Signal Based Surveillance (AVSS), pp.411- 416, Sept 2005.

[4] T.-L. Pao, Y.-T.Chen, J.-H.Yeh, P.-J. Li, "Mandarin emotional speech recognition based on SVM and NN", Proceedings of the 18thInternational Conference on Pattern Recognition (ICPR'06), vol. 1, pp. 1096-1100,September 2006.

[5] Xia Mao, Lijiang Chen, LiqinFu, "Multi-level Speech Emotion Recognition Based on HMM and ANN", 2009 WRI World Congress, Computer Science and Information Engineering, pp.225-229, March 2009.

[6] B. Schuller, G. Rigoll, M. Lang, "Hidden Markov model-based speech emotion recognition", Proceedings of the IEEE ICASSP Conference on Acoustics, Speech and Signal Processing, vol.2, pp. 1- 4, April 2003.

[7] YashpalsingChavhan, M. L. Dhore, PallaviYesaware, "Speech Emotion Recognition Using Support Vector Machine", International Journal of Computer Applications, vol. 1, pp.6-9,February 2010.

[8] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in Proceedings of the 26th Annual Conference on Neural Information Processing Systems (NIPS '12), pp. 1097–1105, Lake Tahoe, Nev, USA, December 2012.

[9] Z. Li, "A study on emotional feature analysis and recognition in speech signal," Journal of China Institute of Communications, vol. 21, no. 10, pp. 18–24, 2000.

[10] T. L. Nwe, S. W. Foo, and L. C. de Silva, "Speech emotion recognition using hidden Markov models,"Speech Communication, vol. 41, no. 4, pp. 603–623, 2003.

[11] X. Bo, "Analysis of mandarin emotional speech database and statistical prosodic features," inProceedings of the Interntional Conference on Affective Computing and Intelligent Interaction (ACLL '03), pp. 221–225, 2003.

[12] L. Zhao, X. Qian, C. Zhou, and Z. Wu, "Study on emotional feature derived from speech signal," Journal of Data Acquistion& Processing, vol. 15, no. 1, pp. 120–123, 2000