

Mood Enhancing Music Player Based on Speech Emotion Recognition and Text Emotion Recognition

Shrikala Madhav Deshmukh¹, Dr. Devulapalli Sita²

¹Research Scholar, Amity School of Engineering & Technology, Amity University, Mumbai, India,
shrikaladeshmukh89@gmail.com

²Professor, Amity School of Engineering & Technology, Amity University, Mumbai, India

ABSTRACT

Music plays a vital role in uplifting mood. A system for “Mood enhancing music player based on speech emotion recognition and text emotion recognition” is proposed in this work. Emotions indicating the mood of a person are derived from speech as well as text input gathered from the person. Speech input is converted to text using speech recognition and Support vector machine (SVM) algorithm is used to find emotion behind speech. Text input is processed using Natural language processing (NLP) and Support vector machine (SVM) algorithm for analyzing emotions in the text. Play list is generated based on a unique mapping from the emotions derived from speech and text inputs to the songs that uplift the mood. A database of songs in Hindi has been prepared for the system and results preliminary experimentation is discussed.

Key words : Natural Language Processing, Speech emotion recognition, Text emotion recognition, Mood Uplifting, Support vector machine

1. INTRODUCTION

Segregation of songs and generation of an appropriate playlist is based on an individual's emotional features is a very tedious and time consuming task. Music has become a part of human life as it plays a key role in expressing feelings. Indian music has a large dataset of songs for every emotion. An attempt is made to categorize Hindi popular music area, under specific moods in our work. Availability of online music libraries has expanded approaches for selecting the songs and has also made it complex as user might get confused or lost in the large dataset. An approach to assist selection based on emotions of the user is proposed in this paper.

Emotion recognition is the process of identifying human emotion. Human emotions – Happy, Sadness, Anger, Fear, Boredom, and Excitement - can be recognized through analysis of Visual/facial expressions, Speech or Text inputs. It is proposed to use Speech and text for emotion recognition for

our work. The task of speech emotion recognition is very challenging for the reasons like variety of sentences, different speakers, various speaking styles, and speaking rates. Another challenging issue is the cultural differences in how a certain emotion is expressed by a speaker. There are mainly two types of speech emotion recognition: a) Speaker Dependent, wherein energy and pitch are used as features for emotion recognition and b) Speaker Independent, in which the focus is on “what was said” regardless “who said it”.

Text emotion recognition can be treated as content-based classification, involving concepts of natural language processing and machine learning. Recognizing and analyzing emotion from text is very challenging task. Detection of human emotions in text is becoming increasingly important for various applications.

We implemented a music player which combines both speech emotion recognition and text emotion recognition. Section 1 discusses background of emotion recognition, speech emotion recognition and Text emotion recognition. Section 2 elaborates proposed system on mood enhancing music player using speech emotion recognition and Text emotion recognition. Section 3 describes database creation of text and songs. Section 4 shows results and analysis. Section 5 describes conclusion.

1.1 Background

For emotion detection, iterative feature normalization (IFN) framework is used by Carlos Busso, et al. [1]. IEMOCAP database has used. Viola-Jones face detection algorithm and KNN classifier is used by Dolly Reney, et al. [2] to detect emotions. DWT, Cepstrum, MFCC and Pitch are used to extract the feature information by B. Rajasekhar, et al. [3]. Telugu database is used consisting of two speakers male and female which contains four emotions namely: Happy, Angry, Sad and Neutral.

A toolkit named as EmoTxt for emotion recognition from text is developed by Calefato F, et al. [4]. Content – based classification is used by Shivhare, et al. [5]. An overview of

emotion detection from text is done by Chopade [6] and emotion detection methods are described.

A music player is implemented by Kabani H et al. [7]. For emotion recognition they had used face emotion recognition system. Emotion based music player is implemented by Hemanth P, et al. [8]. For emotion recognition they had used face emotion recognition system. The emotions are recognized using Support Vector Machine (SVM) algorithm of machine learning.

Support Vector Machine (SVM) algorithm is used by Aaron Don M. Africa et al. [9] for speech emotion recognition as it is one of the leading classifying algorithms in today's time. CNN (Convolutional Neural Network) is used for emotion recognition by Illuri Sreenidhi et al. [10]. They used facial movements to detect emotions.

2. PROPOSED SYSTEM

The proposed system is "Mood enhancing music player based on speech emotion recognition and text emotion recognition", this system generates playlist and play music as per emotion. Input from user is accepted in the form of speech or text and processed using Speech emotion recognition and text emotion recognition.

In speech emotion recognition, Input is the user's speech, when user speaks system converts speech to text. By applying NLP (Natural language processing) algorithms and support vector machine (SVM) algorithm, system finds emotion behind it. For emotion extraction from Text, the music player system displays a text box for the user to input text, NLP (Natural language processing) algorithms and SVM algorithm are used for Text emotion recognition. Based on text provided in the given format, system finds emotion behind provided text.

A database of text has been prepared. A database of songs has been created based on the empirical method of classification by studying the nature of songs and what people listen to for several emotions.

The mood uplifting technique is used for enhancing mood. In mood uplifting technique, motivational songs are used to change feeling/mood of a person from sad/bored to happy. Motivational songs to change mood is based on songs mapped to elevate the mood are used. Here, 6 emotions: Happy, sad, angry, fear, bored and excited are used. The Figure 1 shows the system architecture.

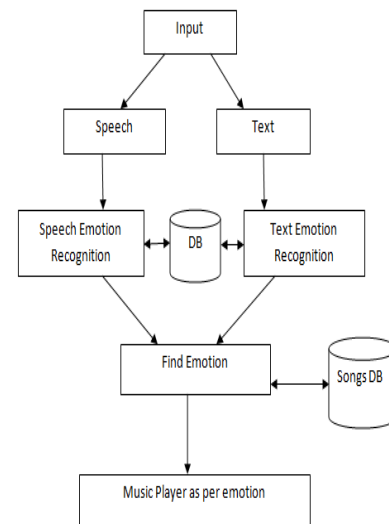


Figure 1: System architecture

3. DATABASE CREATION

Text Database: For text database, Data preprocessing and feature extraction techniques are used. For Data preprocessing, The Natural Language Toolkit is one of the best-known and most-used NLP libraries is used. For feature extraction, Bag of Words (BOW) and word embedding techniques are used. Training and Testing are applied on dataset. The Table 1 shows the used text database.

Table 1: Text database

Emotion	Words
Happy	Adore, agreeable, appealing, beautiful, bliss, blissful, carefree, charming, cheerful, favorable, fine, friendly, fulfilled, fun, gladden, glorious, glory, glory in, good, grateful, great, happy, happiness, joy, lucky, nice, etc.
Sad	Lonely, heartbroken, gloomy, disappointed, hopeless, grieved, unhappy, lost, troubled, resigned, miserable, etc.
Angry	Annoyed, frustrated, peeved, contrary, bitter, infuriated, irritated, mad, cheated, vengeful, insulted, etc.
Fear	Worried, doubtful, nervous, anxious, terrified, panicked, horrified, desperate, confused, stressed
Bored	Disgust, fatigue, monotony, detachment, distaste, dullness, unconcern, lack of interest, etc.
Excited	Energetic, playful, cheerful, joyful, amazed, etc.

Songs Database: A database of Hindi songs for all emotions has been created. It consists of 100 songs per emotion. For Training and Testing of songs database, HMM algorithm is used. In addition to this, mood uplifting technique is used. As we know, music is unquestionably important in shaping moods. So, if someone is sad/ bored, then instead of playing sad/boring songs, system is playing motivational music such that his/her mood can change. The Table 2 shows the used songs database.

Table 2: Songs database

Emotion	Songs
Happy	Track 01 “Roobaroo” Track 02 “Sanorita” Track 03 “Illahi”, etc...
Sad (Motivational)	Track 01 “Rook jana nhi” Track 02 “Apna time aayega” Track 03 “All is well”, etc..
Angry	Track 01 “Sarkar” Track 02 “Atyachar” Track 03 “Gusse se”
Fear	Track 01 “Aankhen milayenge” Track 02 “Darna mana hai” Track 03 “Darr”
Bored (Motivational)	Track 01 “Life sahi hai” Track 02 “Jee le jara” Track 03 “Tu hai sikandar”
Excited	Track 01 “Rock on” Track 02 “Khwaboon ke parinde” Track 03 “Salaam India”

4. RESULTS AND ANALYSIS

4.1 Accuracy of Emotion Recognition

Natural language processing (NLP) and Support vector machine (SVM) algorithm are used for Speech emotion recognition (SER) and Text emotion recognition (TER). Result analyzed by testing the system 100 times for each emotion: happy, sad, angry, fear, bored and excited. Here are some preliminary results of system using Natural language processing (NLP) and Support vector machine (SVM) algorithm. This system is showing 97% accuracy for happy emotion recognition. The Figure 2 shows accuracy of emotion recognition.

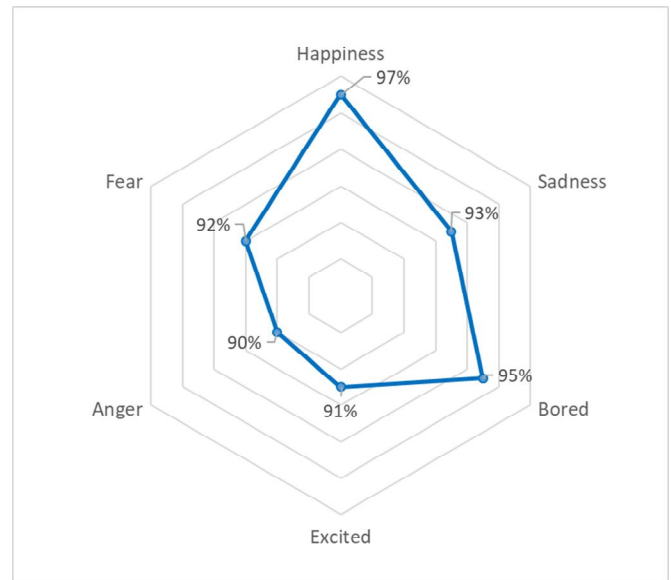


Figure 2: Accuracy of emotion recognition

4.2 Accuracy of Music Player

The Figure 3 shows the accuracy of music player.

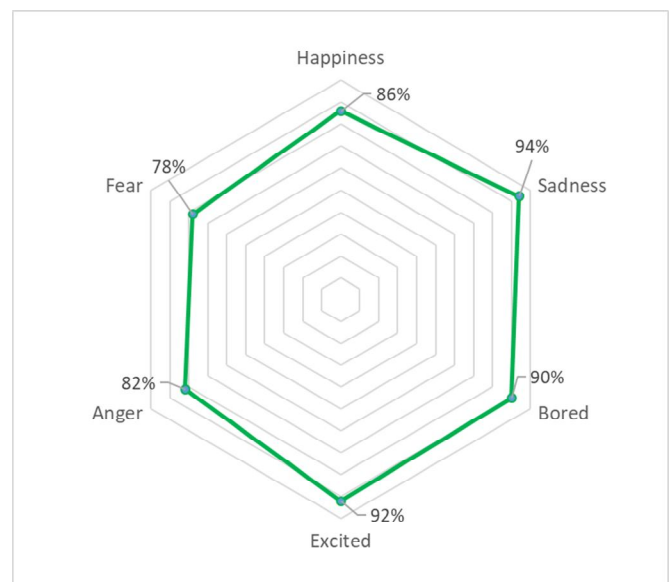


Figure 3: Accuracy of music player

5. CONCLUSION

In this paper, a system on “Mood enhancing music player based on speech emotion recognition and text emotion recognition” is implemented. Here, playing a song based on the person’s emotions is implemented. In the system, input is of two types: speech and text. In speech emotion recognition, input is user’s speech, then system will convert speech to text and after applying NLP and SVM algorithm, system finds emotion behind speech. In text emotion recognition, textual data is input to the system. After applying NLP and SVM algorithm, system finds emotion behind text. Six emotions are

selected: happy, sad, angry, fear, bored and excited. Output is a song based on the emotion detected from the speech or text. Mood uplifting technique is used for enhancing mood.

REFERENCES

1. C. Busso, S. Mariooryad, A. Metallinou, S. Narayanan. **Iterative Feature Normalization Scheme for Automatic Emotion Detection from Speech**, *IEEE Transactions on Affective Computing*, pp. 1-14, 2012.
2. D. Reney, N. Tripaathi. **An Efficient Method to Face and Emotion Detection**, *Fifth International Conference on Communication Systems and Network Technologies, IEEE*, 2015, pp. 493-497, 2015.
<https://doi.org/10.1109/CSNT.2015.155>
3. B. Rajasekhar, M. Kamaraju and V. Sumalatha. **Speaker Independent Emotion Recognition from Speech Signals**, *ARN Journal of Engineering and Applied Sciences*, vol. 13, no. 9, pp. 3188-3191, 2018.
4. F. Calefato, F. Lanubile, N. Novielli. **EmoTxt: a toolkit for emotion recognition from text**. In *2017 seventh international conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW)*, *IEEE*, pp. 79-80, 2017.
5. S. N. Shivhare, S. Khethawat. **Emotion Detection From Text**, *CS & IT-CSCP*, pp. 371–377, 2012.
<https://doi.org/10.5121/csit.2012.2237>
6. C.R. Chopade. **Text based emotion recognition: A survey**. *International journal of science and research*, vol. 4, no. 6, pp. 409-414, 2015.
7. H. Kabani, S. Khan, O. Khan and S. Tadv. **Emotion based music player**, *International Journal of Engineering Research and General Science*, 3(1), pp. 2091-2730, 2015.
8. P. Hemanth, A.C. Adarsh, P.B. Ajith, V. Kumar. **EMO PLAYER: Emotion Based Music Player**, *International Research Journal of Engineering and Technology (IRJET)*, vol. 5, no. 4, pp. 4822-4827, 2018.
9. Aaron Don M. Africa, Anna Rovia V. Tabalan, Mharela Angela A. Tan. **Speech Emotion Recognition using Support Vector Machines**, *International Journal of Emerging Trends in Engineering Research (IJETER)*, vol 8. no. 4, pp. 1212-1216, April 2020.
<https://doi.org/10.30534/ijeter/2020/43842020>
10. Illuri Sreenidhi, P. Satyanarayana. **Real-Time Human Fall Detection and Emotion Recognition using Embedded Device and Deep Learning**, *International Journal of Emerging Trends in Engineering Research (IJETER)*, vol 8. no. 3, pp. 780-786, March 2020.
<https://doi.org/10.30534/ijeter/2020/28832020>