# An Emotion Detection System Based on Multi Least Squares Twin Support Vector Machine

[1]Mr.S.Prabu, [2]Ms.M.Vinothini, [3]Ms.P.Vinithasri, [4]& Ms.C.Suganya,

[1]Assistant Professor, CSE, Karpaga Vinayaga College Of Engg & Tech
[2,3,4]IV Year, CSE, Karpaga Vinayaga College Of Engg & Tech

[1]sprabumkm@gmail.com
[2]mvino95@gmail.com

***Abstract*--Posttraumatic stress disorder (PTSD), bipolar manic disorder (BMD), obsessive compulsive disorder (OCD), depression, and suicide are some major problems existing in civilian and military life. The change in emotion is responsible for such type of diseases. So, it is essential to develop a robust and reliable emotion detection system which is suitable for real world applications. Detection of emotion in speech can be applied in a variety of situations to allocate limited human resources to clients with the highest levels of distress or need, such as in automated call centers or in a nursing home. In this paper, we used a novel multi least squares twin support vector machine classifier in order to detect seven different emotions such as anger, happiness, sadness, anxiety, disgust, panic, and neutral emotions. The experimental result indicates better performance of the proposed technique over other existing approaches. The result suggests that the proposed emotion detection system may be used for screening of mental status.**

## I. Introduction

Stressful situation can cause some major psychiatric problems such as depression, suicide, PTSD, BMD, and OCD in civilian as well as in military life. Earlier treatment may become useful for such type of psychiatric problems [1]. So, there is a need to develop technology for recognizing early change in human behavior. Several biomarkers are reported by the medical researchers for psychiatric diseases. But these biomarkers are not effective in military life as they required a big and complicated machine for detecting psychiatric diseases. On the other hand, there is a fast development in voice, speech, and emotion detection technologies in engineering field. These technologies provide human-machine interaction for emotion detection and further treatment of psychiatric problems. Several researches measured the level of fatigue and stress from speech. But the level of fatigue and stress does not lead to psychiatric disorder directly. Emotion change of a human can cause mental diseases. Mostly, clinicians recognize the mental state of a patient from his/her face and voice which represents his/her emotion. This fact leads to the possibility that emotion detection system can be used for recognizing the mental disorder or disease in human. Early detection of disease improves the prognosis and is helpful to provide effective treatment at early stages. Emotion detection system can provide support to the clinicians to perform the task of emotion detection more efficiently. In automated call centers or in a nursing home, while nursing staff may not be available to assist everyone, automated emotion detection can be used to “triage” a patient. Automated emotion detection system is helpful to recognize whether a patient becomes angry or impatient and if so then the staff or treatment is provided to that patient as soon as possible.

Nowadays emotion detection from speech is an active research area and is useful for man-machine interaction. Various researches have been done about automated emotion detection from facial expressions. But this task is computationally expensive and complex due to the requirement of high quality cameras for capturing face images. Apart from facial expression, emotions are also detected from speech which has been proven to be more promising modality. Since speech is the primary mode of human communication, the detection of emotion from speech is an important aspect.

Machine learning algorithms such as -nearest neighbor (NN), artificial neural network (ANN), and support vector machine (SVM) are widely used for emotion detection due to their excellent performance [8–13]. In this paper, the proposed emotion detection system recognizes seven different emotions which are anger, anxiety, disgust, happiness, sadness, panic, and neutral emotions. Different emotions can be seen as different classes. So, it requires a multiclassifier for emotion detection. In this paper, we proposed a novel multi least squares twin support vector machine (MLSTSVM) classifier which is the extension of binary least squares twin support vector machine (LSTSVM). So, the proposed system predicts the class or emotion for a given input. In order to check the validity of the proposed classifier, we evaluated its performance against 5 benchmark datasets.

The paper is organized as follows: introduction section includes need for emotion detection system. Section 2 pro-vides the detail of our novel classifier which is multi least squares twin support vector machine. Proposed framework for emotion detection and dataset details are discussed in Section 3. The experimental results and conclusion of the proposed emotion detection system are presented in Sections 4 and 5, respectively.

## II. Multi Least Squares Twin Support Vector Machine

LSTSVM for binary classification which solves two linear programming problems and constructs two nonparallel hyper planes, one for each class. Since real world data contains multiple classes and requires a classifier that works well for multiple classes, in this paper, we propose a novel multiclassifier termed as MLSTSVM. This classifier is an extension of the binary LSTSVM and is based on "one-versus-rest" strategy. Here, we selected and extended the binary LSTSVM because it shows better generalization ability and is faster as compared to other existing approaches [14, 15]. MLSTSVM constructs " " hyper planes, one for each class, by optimizing -linear programming problems, where " " denotes number of classes. It adopts the concept of "one-versus-rest" in which the data points of each class are trained with the data points of other classes. Consider dataset has " " number of data points in training dataset: { $_1$, $_1$},...,{ , }. Here is a feature vector in -dimensional space and ∈ {1,2,..., }is the label of corresponding class. "One-versus-rest" generates binary LSTSVM classifier, each of which separates one class from the rest of the classes. The LSTSVM classifier assumes the data points of the class as positive data points and the data points of other classes as negative data points. Advances in Artificial Intelligence The matrix includes all the data points except the class. MLSTSVM classifier for both linear and nonlinear cases is formulated as follows.

***2.1. Linear Case***. The equation of the hyperplane is obtained as term, respectively, in real space . The LSTSVM classifier slack variable correspondingly. The first term of (3) denotes the squared sum distance of the data points of the class. The minimization of this term keeps the hyperplane in the close affinity of the class. The second term of (3) minimizes the misclassification error of the data points of rest of the -1 classes. So, in this way the hyperplane is kept in the close affinity with the data points of th class and lies as far as possible from the data points of other classes. The objective function is solved by taking its dual form. Lagrangian function of the objective function as mentioned where > 0 represents the Lagrangian multiplier. The optimization of Lagrangian function is achieved by differentiating it with respect to normal vector, bias, slack variable, and Lagrangian multiplier and the following Karush-Kuhn-Tucker (KKT) conditions are obtained:

## III. Description of Dataset and Proposed Model

***3.1. Dataset Description.*** Emotion detection system from human speech is divided into two parts: dataset collection and feature extraction system and emotion detection system as shown in Figure 1. Emotion dataset, used in this research work, is taken from two sources: a benchmark dataset from source http://www.eecs.qmul.ac.uk /mmv/datasets/ deap/download.html and real dataset (audio file) collected using Nexus 5 smart phone. The benchmark dataset contains the voice recordings of expert reciting numbers and dates with different emotions and real dataset contains the audio recording of 10 male persons where each person voice is recorded minimum 2-3 times with different sets of emotions.

The dataset includes seven different emotions such as anger, anxiety, happiness, sadness, disgust, neutral emotions, and panic for emotion detection. Audio file of a person or a patient is given for emotion detection. Power Audio Cutter is used to cut the audio files and the duration of each audio file is 2 seconds. This research used PRAAT scripting tool for feature extraction which is a freeware and flexible tool developed by Paul Boresma and David Weenink of the University of Amsterdam for speech analysis. It performs spectral, formant, pitch, intensity, jitter, and shimmer analyses

This tool generates voice report from audio files and converts the audio files into text files. The voice report of these audio files contains different features of voice like pitch, intensity, shimmer, jitter, and so forth. Pitch, also known as vibration rate of the vocal folds, is one of the most important and essential parts of the human voice. The sound of the voice varies according to the vibration rate. High pitch refers to high vibration rate which further increases the sound of the voice while low pitch corresponds to the lower sound. Vibration rate is

Dependent on the duration and thickness of vocal cords. Relaxation and tightening of Intelligence the muscles around vocal cords also affect the vibration rate. Emotion or mood of a person also has an effect on his/her pitch. During excitement or fright, the muscles put strain on vocal cords which further produce high pitch voice. The tone of a person describes the way a statement is presented and can convey the emotion, psychological arousal, and mood of that person. Usually, softer pitch and tone are seen as nonaggressive and indicate the friendly behavior of a person. Jitter and shimmer are another important attribute of a voice. Jitter and shimmer measure the irregularity percentage in the pitch and in the amplitude of the vocal note correspondingly. Voice quality and signal-to-noise ratio can be estimated from harmonicity. The second part includes significant feature selection from voice report generated by PRAAT and classification of emotions using MLSTSVM. Feature selection (FS) is the process of selecting relevant and important attributes from a dataset and plays a significant role in the construction of a classification system [19–21]. FS is also termed as attribute selection process which reduces the number of input attributes by selecting only important attributes for a classifier in order to enhance its performance. In this paper, we used the combination of -score and sequential forward selection approaches for feature selection.

## IV. Experimental Results and Discussion

The benchmark and real data exist in the form of audio files. The experiment is performed on the time span of 2 seconds for each audio file for which a tool Power

Audio Cutter is used to cut the audio files as per required duration. The feature of audio files is extracted by using PRAAT scripting tool. Figure 3 shows the browsing of audio file for running PRAAT script in order to extract features from it. The voice report of an audio file generated by using PRAAT.

*Table 1: Dataset details.*

| Emotions | Number of instances | Number of attributes |
|---|---|---|
| Anger | 54 | 24 |
| Anxiety | 30 | 24 |
| Happiness | 60 | 24 |
| Sadness | 46 | 24 |
| Disgust | 36 | 24 |
| Neutral emotions | 32 | 24 |
| Panic | 32 | 24 |

*Table 4: Average value of -score using 10-fold cross validation.*

| Features | -score | Features | -score |
|---|---|---|---|
| F1 | 3.6977 | F13 | 0.0090 |
| F2 | 1.6892 | F14 | 0.1872 |
| F3 | 0.0966 | F15 | 0.0074 |
| F4 | 0.2402 | F16 | 0.0088 |
| F5 | 0.1119 | F17 | 0.0171 |
| F6 | 0.1250 | F18 | 0.0070 |
| F7 | 0.1143 | F19 | 0.0108 |
| F8 | 0.0187 | F20 | 0.0163 |
| F9 | 0.0095 | F21 | 0.0012 |
| F10 | 0.0588 | F22 | 0.0039 |
| F11 | 0.0106 | F23 | 0.0168 |
| F12 | 0.0405 | F24 | 0.0010 |

The above table shows the snapshot of the emotion detection dataset. In this snapshot 35 instances, 5 instances of each emotion, have been taken to generate a complete view of range of various attributes for corresponding class. The first attribute of the snapshot denotes emotions. Here, "1," "2," "3," "4," "5," "6," and "7" are used for anger, anxiety, disgust, happiness, neutral emotions, panic, and sadness, respectively.

Since the range of attributes varies from each other, normalization of each attribute value is performed to take them within the specified range. Two feature selection techniques, -score and SFS, are used for selecting significant features. Table 4 shows the average value of -score for each attribute or feature by using 10-fold cross validation. After calculating the -score of each attribute, SFS is used for obtaining 24 feature subsets or models. The importance or -score values of each feature from high to low are 1, 2, 4, 14, 6, 7, 5, 3, 10, 12, 8, 17, 23, 20, 19, 11, 9, 13, 16, 15, 18, 22, 21, and 24. Table 5 shows the twenty-four feature subsets or models on the basis of SFS. For each feature subset, a MLSTSVM classifier is constructed and its predictive accuracy is checked using 10-fold cross validation method. The proposed MLSTSVM classifier is implemented using MATLAB R2012a.

## V. Conclusion

The proposed emotion detection system can be used in automated call centers or in a nursing home where resources or nursing staff may not be available to aid everyone. Auto-mated emotion detection system can be useful to identify the emotion change of patients and to trigger the alarm according to their emotion change so that effective treatment or facility can be provided to patients as soon as possible. This system can assist the clinician to perform the task of emotion detection more efficiently. The proposed emotion detection system may serve as an important tool because change in emotion is responsible for several diseases such as PTSD, BMD, OCD, depression, and suicide. In this paper, emotions are detected by using a novel classifier, named MLSTSVM, and its performance is validated on five benchmark datasets. PRAAT scripting tool is used for feature extraction and extracted 24 features from voice recording. The combination of -score and SFS is used for selecting important features from emotion detection dataset. It is found that MLSTSVM classifier based emotion detection system with sixteen features has achieved better predictive accuracy, 87.28% for linear MLSTSVM, 92.89% for Gaussian MLSTSVM, and 88.87% for polynomial MLSTSVM classifier. T he performance of proposed system is compared with NN, ANN, and multi-SVM approaches. Experimental results indicate that our proposed novel classifier based emotion detection performs well as compared to the other existing approaches. The results of proposed classifier are also verified by using real dataset containing the voice of 10 persons with different emotions and obtained 86.18% accuracy with Gaussian MLSTSVM. The whole system may be adopted and extended as an intelligent personal assistant application for helping disable, autistic children, psychic patients, and elderly people. Apart from healthcare, importance of automatically recognizing emotions from human speech as achieved here in this proposed system may also be used as a part of human computer interaction applications such as robotics, games, and intelligent tutoring system. We have developed the emotion recognition system using MATLAB on Windows operating system. The system has certain limitations; for example, it does not deal with the background noise and is trained for male voices only. Hence few enhancements are possible in the future. A better performance could be guaranteed by optimizing the values of certain parameters like sigma (for Gaussian kernel function) and cost parameters by using genetic algorithm, particle swarm optimization, or any other optimization approaches.

## References

*[1] S. Tokuno, G. Tsumatori, S. Shono et al., "Usage of emotion recognition in military health care," in Proceedings of the Defense Science Research*

Conference and Expo (DSR '11), pp. 1–5, Singapore, August 2011.

[2] A. H. Miller, V. Maletic, and C. L. Raison, "Inf lammation and its discontents: the role of cytokines in the pathophysiology of major depression," Biological Psychiatry, vol. 65, no. 9, pp. 732– 741, 2009.

[3] C.-N. Anagnostopoulos, T. Iliou, and I. Giannoukos, "Features and classifiers for emotion recognition from speech: a survey from 2000 to 2011," Artificial Intelligence Review, 2012.

[4] S. G. Koolagudi and K. S. Rao, "Emotion recognition from speech: a review," International Journal of Speech Technology, vol. 15, no. 2, pp. 99–117, 2012.

[5] S. Ramakrishnan and I. M. M. El Emary, "Speech emo-tion recognition approaches in human computer interaction," Telecommunication Systems, vol. 52, no. 3, pp. 1467–1478, 2013. K. Shiomi, "Voice processing technique for human cerebral activity measurement," in Proceedings of the IEEE International Conference on Systems, Man and Cybernetics (SMC '08), pp. 3343–3347, October 2008.

[6] T. Pfister, "Emotion Detection from Speech," 2010.

[7] H. Sato, Y. Mitsukura, M. Fukumi, and N. Akamatsu, "Emo-tional speech classification with prosodic prameters by using neural networks," in Proceedings of the 7th Australian and New Zealand Intelligent Information Systems Conference, pp. 395– 398, IEEE, 2001.

[8] M. El Ayadi, M. S. Kamel, and F. Karray, "Survey on speech emotion recognition: features, classification schemes, and databases," Pattern Recognition, vol. 44, no. 3, pp. 572–587, 2011.

[9] H. Altun and G. Polat, "Boosting selection of speech related fea-tures to improve performance of multi-class SVMs in emotion detection," Expert Systems with Applications, vol. 36, no. 4, pp. 8197–8203, 2009.

[10] D. Tomar and S. Agarwal, "A survey on data mining approaches for healthcare," International Journal of Bio-Science and Bio-Technology, vol. 5, no. 5, pp. 241–266, 2013.

[11] D.-Y. Huang, Z. Zhang, and S. S. Ge, "Speaker state classif ication based on fusion of asymmetric simple partial least squares (SIMPLS) and support vector machines," Computer Speech and Language, vol. 28, no. 2, pp. 392–414, 2014.