

<http://ansinet.com/itj>

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSI*net*

Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Comparison of Performance Using GA and FIS in Spotting the Aberration Spot of Clinical Data with the Aid of NN and SVM

¹C.R. Bharathi and ²V. Shanthi

¹Research Scholar, Sathyabama University, AP, VelTech University, Chennai, India

²Department of MCA, St. Joseph's College of Engineering, Chennai, India

Abstract: The main concern of the present investigation is to compare the performance between (1) Neural Network (NN) and Support Vector Machine (SVM) and (2) Genetic Algorithm (GA) and Fuzzy Inference System (FIS) which were used to identify the specific (acute) spots of the abnormal children speech who need specific practice in those identified spots to overcome stuttering behavior. In the previous work, there are three phases followed (1) Training Phase (2) Testing phase and (3) Optimization phase for blemishing the acute spot in which the speech practice can be provided for the abnormal children. The training and optimization in previous works were implemented using (1) NN and GA and (2) SVM and FIS. In this study, the comparison is done between the performance of (1) NN and SVM and (2) GA and FIS using Mathews Correlation Coefficient (MCC). Performance of the proposed system is tested by the dataset generated by abnormal and normal female children. The proposed system is implemented using MATLAB as working platform. This study is to provide a system to identify the acute spots in abnormal word where the stuttering children need the practice. Comparison Study is done between the performance of (1) NN and SVM and (2) GA and FIS.

Key words: Speech signal, stress, mel frequency cepstrum coefficients, principal component analysis, support vector machines, fuzzy inference system, genetic algorithm

INTRODUCTION

In real-world environment, speech signal processing plays a vital role among the research communities in general and it has a specific role in the case of children who have the problem of stuttering in particular. But, experts feel that the latter case is caused by a combination of various factors. First, genetics is believed to play a part. It appears that stuttering results from inherited (genetic) abnormalities in the language centres of the brain. Second, children who have developmental delays or other speech problems are more likely to stutter. Third, environmental factors can have an influence. Speech synthesis among the children helps in identifying the acute spot in the speech of abnormal children.

In this study, the abnormal children speech is considered to enhance the speech by comparing it with the normal children speech. This work is very useful for the speech practitioners to improve the speech of the abnormal children by knowing where the speech practice is required. The Mel Frequency Cepstrum Coefficients (MFCC) for abnormal and normal speeches was extracted. Principal Components Analysis (PCA) is applied to reduce the dimensionality of MFCC features and is trained

with the aid of Neural Network (NN) and Support Vector Machine (SVM) separately to identify the abnormal word (Bharathi and Shanthi, 2012c). Through the thresholding operation acute word is identified. Number of peaks and the amplitude is found from the Fast Fourier Transform (FFT) of the acute word. The amplitude and number of peaks are converted into genes and optimized using Genetic Algorithm (GA). And also the amplitude and number of peaks are inputted to the Fuzzy Inference System (FIS) for the optimization process separately to identify the position in which the stress has to given. The comparison of performance between (1) NN and SVM and (2) GA and FIS is done in this study.

MATERIALS AND METHODS

Classification process

Classification through NN: In the classification, the NN is utilized for the classification purpose in order to separate the normal and for identifying the abnormal dataset. The seven MFCC features namely mean, standard deviation, maximum amplitude value and its id, minimum amplitude value and its id, MFCC length were extracted and the same set of seven features were extracted from the

original word. These fourteen features are trained using NN and SVM separately (Bharathi and Shanthi, 2011a; Bharathi and Shanthi, 2012c). This neural network has fourteen input layers which are the parameters of a word. Hidden layers and one output layer to identify the word inputted is either normal or abnormal.

Support Vector Machine (SVM): An SVM is a concept in statistics and computer science for a set of related supervised learning methods that analyze data and recognize patterns, used for classification and regression analysis. The standard SVM takes a set of input data and predicts, for each given input, which of two possible classes forms the input, making the SVM a non-probabilistic binary linear classifier. Given a set of training examples, each marked as belonging to one of two categories, an SVM training algorithm builds a model that assigns new examples into one category or the other.

Identifying the acute spot of aberration of abnormal children using GA and FIS

Genetic Algorithm (GA): In Optimization Phase, thresholding operation on the identified abnormal word is done and then FFT is computed for the acute word. The sudden changes in frequency can be seen in FFT by maximum peaks (Bharathi and Shanthi, 2011b). The numbers of maximum peaks are analyzed with its amplitude using FFT. For identifying the acute spot of aberration for abnormal children, GA is used. This each peak is considered as a gene. Number of Chromosomes obtained from crossover and mutations are combined and so population pool is filled up. After crossover process, the fitness is applied on children chromosomes and hence highly deviated chromosomes are replaced with new genes. The process is repeated until it reaches maximum iteration. The indices, which are obtained from genes of best chromosomes, represent the acute position of abnormal word (Bharathi and Shanthi, 2012e).

Fuzzy Inference System (FIS): Fuzzy knowledge-based systems are rules that built on fuzzy logic and fuzzy set theory. A fuzzy rule system is a rule system whose variables or part of its variables are linguistic variables. It is well known that FIS can be used to approximate closely any nonlinear input output mapping by means of a series of if-then rules. In the design of FIS, there are two major tasks, viz. the structure identification and the parameter adjustment. The fuzzy inference system is comprised of three phases that are Fuzzification, Rules Evaluation and Defuzzification. The process of fuzzy inference involves Membership Functions, Logical Operations and If-Then Rules. In this FIS, the peaks of the abnormal words are utilized in order to identify the aberration spot (Bharathi and Shanthi, 2012a).

Experimental results: There are number of approaches and among them few approaches are detailed below (Bharathi and Shanthi, 2012d):

- Automatic detection method which uses a new approach called the speech recognition technology for syllable repetition in read speech for objective assessment of stuttered disfluencies. Different steps were involved in finding the number of repetitions from the speech samples using MFCC feature extraction algorithm. The twelve dimensional MFCC obtained for each syllable are used to compute the angle between them which serve as local-distance and represent in the form of matrix. Using Dynamic Programming the minimum-cost path through matrix is found and given to Decision logic. In Decision logic, Dynamic Time Warping (DTW) based score matching is done
- Intelligent processing of stuttered speech work reports results on automatic detection of speech disorder events that were based on both rough sets and artificial neural networks. Rough set-based algorithm and Neural Network algorithm was applied in (a) Automatic detection of stops-gaps, (b) Discerning vowel prolongations and (c) Detection of syllable repetitions. During the tests, rough set system settings were selected as follows: the value of quantization was from the range of 5 to 20 and the value of neutral point was from 0.6 to 0.9. Better scores were obtained using rough set-based system
- In Optimal Curve Fitting method the amplitude profile of sampled speech data were fitted by employing sum of sine functions with a confidence level. Furthermore, amplitude correlation technique is applied between original speech signal samples of normal and pathological subjects and correlation technique is also applied between the curve fit constant values for normal and pathological subjects. For the collected sample space of children the curve fitting technique is applied and curve fitting constant values are obtained. Then correlation coefficient between normal and pathological curve fitting constant values are calculated and it is observed that if correlation coefficient value starts deviating from its expected value 0.20, speech disability severity increases in pathological subjects in comparison with normal child. In the present, investigation the scope of the proposed work is detailed below to overcome shortcomings observed in the above existing methods

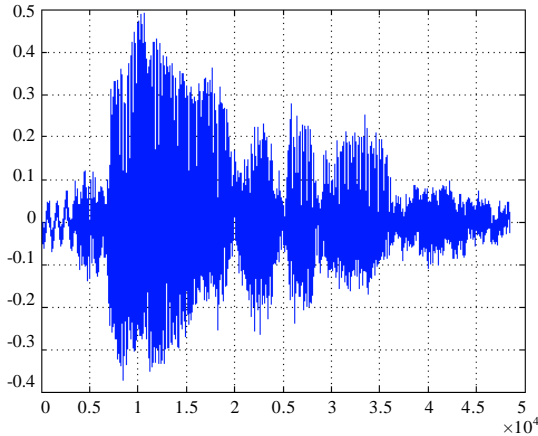


Fig. 1: Signal for the word Dinosaurs (Normal Child1)

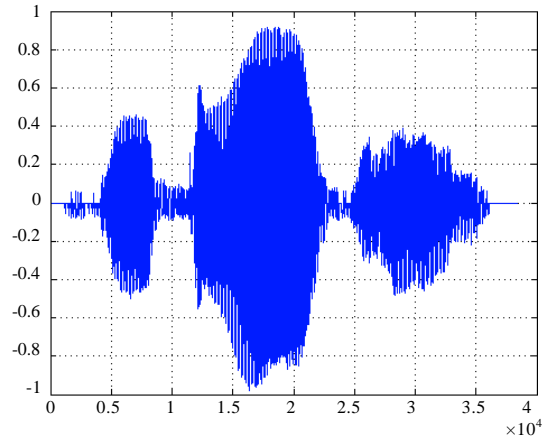


Fig. 3: Signal for the word Dinosaurs (Abnormal Child)

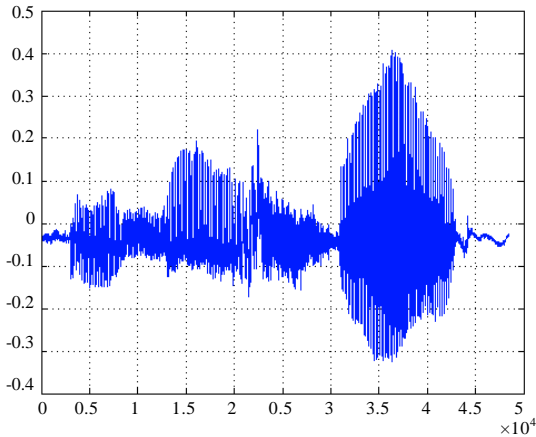


Fig. 2: Signal for the word Dinosaurs (Normal Child2)

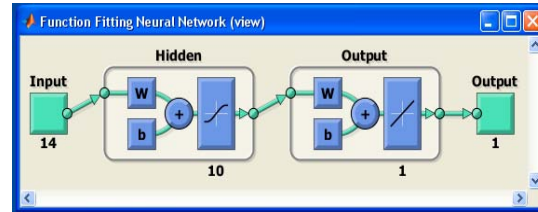


Fig. 4: Generated ANN for classification

Feature extraction: Initially, the words are extracted from the both normal and abnormal children and then the MFCC feature has been extracted from it. Subsequently, the PCA is applied to reduce the dimensionality of the words. Few speech samples which are sent as input to MFCC Feature Extraction (Bharath and Shanthi, 2011c).

Discriminant analysis using SPSS model was implemented on fourteen extracted features and around 80% classification accuracy was got (Bharathi and Shanthi, 2012b).

Classification using NN and SVM: Subsequently, for MFCC features the PCA is applied to reduce the dimensionality of the words and then they are inputted to the NN and SVM separately to identify the abnormal and the normal word.

Figure 4 shows the output generated by NN and Fig. 5 shows the regression result of the SVM Classifier.

The SVM regression results of the recorded dataset of normal and abnormal words are shown in Fig. 5. Signal for the abnormal word Dinosaurs is shown in Fig. 3. Signal for the normal word Dinosaurs of the normal child1 is shown in Fig. 1. Signal for the normal word Dinosaurs of the normal child2 is shown in Fig. 2. Table 1 shows the output for the SVM classifier for the identification of abnormal word. The output results identified by the SVM and NN classifiers are manually verified and classification accuracy is found. The output of SVM classifier and NN classifier for few words as an sample are mentioned in the Table 2 for identification of abnormal and normal word. Ninty percent of Classification accuracy is got by SVM classifier and 60% of classification accuracy is got by NN.

The below rules were also used to find the performance the NN and SVM:

True positive (TP)-Abnormal word identified as abnormal
True Negative (TN)-Normal word identified as abnormal
False positive (FP)-Normal word identified as normal
False Negative (FN)-Abnormal word identified as normal

This performance is verified using the TP, TN, FP and FN rules. Sensitivity and specificity are statistical

Table 1: Few Output for the SVM classifier for the identification of abnormal word

Original	Proposed (SVM)
grapes.wav	Abnormal word
kumudha.wav	Abnormal word
Sugar	Abnormal word
Fish	Normal word
Dinosaur	Abnormal word
ink pot	Abnormal word
Umbrella	Abnormal word
Rabit	Abnormal word
Wild animals	Normal word
Aeroplane	Abnormal word

Table 2: Output for the SVM classifier and ANN classifier for the identification of abnormal word

Original	Proposed (SVM)	Existing (ANN)
analysis.wav	Abnormal word	Normal word
apple.wav	Abnormal word	Normal word
Dinosaurs	Abnormal word	Normal word
Fish	Abnormal word	Abnormal word
Flower	Abnormal word	Abnormal word
ink pot	Abnormal word	Normal word
Umbrella	Abnormal word	Abnormal word
Uniform	Abnormal word	Abnormal word
Utensil	Normal word	Abnormal word
Wild animals	Abnormal word	Abnormal word
Classification accuracy	90%	60%

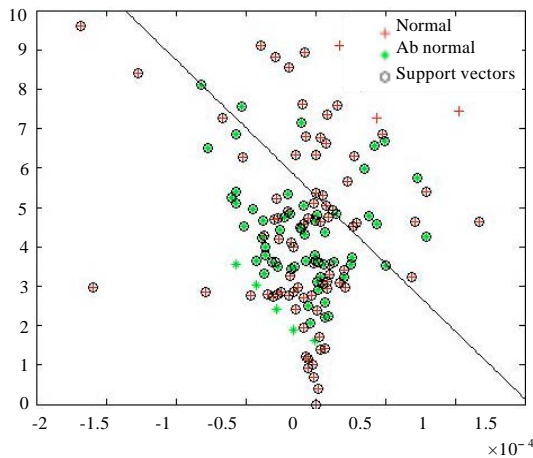


Fig. 5: Regression result for the SVM classification to identify the abnormal word

measures of the performance of a binary classification test, also known in statistics as classification function. Sensitivity also called the true positive rate, measures the proportion of actual positives which are correctly identified as such. Specificity measures the proportion of negatives which are correctly identified as such. Equations 1 to 7 are used for this statistical measures:

$$\text{Sensitivity} = \frac{\text{No. of true positives}}{\text{No. of true positives} + \text{No. of false negatives}} \quad (1)$$

Table 3: Performance of Neural Network for identifying the abnormal words

TP	TN	FP	FN	Sensitivity	FPR	Accuracy	Specificity	PPV	NPV	FDR	MCC
3	2	3	2	60	60	50	40	50	50	50	0.00
3	2	3	2	60	60	50	40	50	50	50	0.00
3	3	2	2	60	40	60	60	60	60	40	0.00
4	2	2	2	67	50	60	50	67	50	33	0.41

Table 4: Performance of Support Vector Machine for identifying the abnormal words

TP	TN	FP	FN	Sensitivity	FPR	Accuracy	Specificity	PPV	NPV	FDR	MCC
4	2	2	2	67	50	60	50	67	50	33	0.41
5	2	2	1	83	50	70	50	71	67	29	0.81
5	3	2	0	100	40	80	75	71	100	29	0.82
6	2	2	0	100	50	80	50	75	100	25	1.00

$$\text{Specificity} = \frac{\text{No. of true negatives}}{\text{No. of true negatives} + \text{No. of false positives}} \quad (2)$$

$$\text{False Positive Rate (FPR)} = \frac{\text{FP}}{\text{N}} = \frac{\text{FP}}{(\text{FP} + \text{TN})} \quad (3)$$

$$\text{Positive Predictive Value (PPV)} = \frac{\text{TP}}{(\text{TP} + \text{FP})} \quad (4)$$

$$\text{Negative Predictive Value (NPV)} = \frac{\text{TN}}{(\text{TN} + \text{FN})} \quad (5)$$

$$\text{False Discovery Rate (FDR)} = \frac{\text{FP}}{(\text{FP} + \text{TP})} \quad (6)$$

$$\text{Mathews Correlation Coefficient (MCC)} = \frac{(\text{TP} \cdot \text{TN} - \text{FP} \cdot \text{FN})}{\sqrt{\text{FP} \cdot \text{FN} \cdot \text{TP} \cdot \text{TN}}} \quad (7)$$

Accuracy (Acc) is calculated as:

$$\text{Acc} = (\text{TP} + \text{TN}) / (\text{TP} + \text{TN} + \text{FP} + \text{FN})$$

The Matthews Correlation Coefficient (MCC) is a measure of the quality of binary (two-class) classifications. It takes into account true and false positives and negatives and is generally regarded as a balanced measure which can be used even if the classes are of very different sizes. The MCC is in essence a correlation coefficient between the observed and predicted binary classifications; it returns a value between -1 and +1. A coefficient of +1 represents a perfect prediction, 0 no better than random prediction and -1 indicates total disagreement between prediction and observation.

The accuracy comparison of different words using NN and SVM is shown in Fig. 6 using the accuracy found in Table 3 and 4 which shows the performance of NN and SVM respectively. Four different groups of ten different words were considered for the Table 3 and 4 to calculate the accuracy and other parameters mentioned in Table 3 and 4.

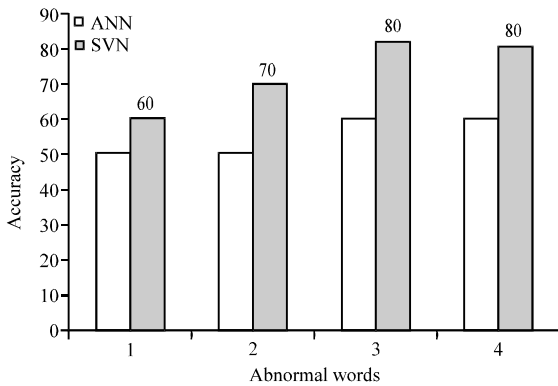


Fig. 6: Accuracy comparisons of different words in ANN and SVM

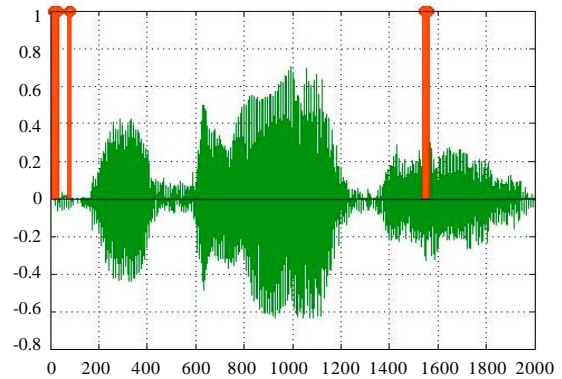


Fig. 8: Identified acute spot of aberration for the word Dinosaurs using FIS (Abnormal child)

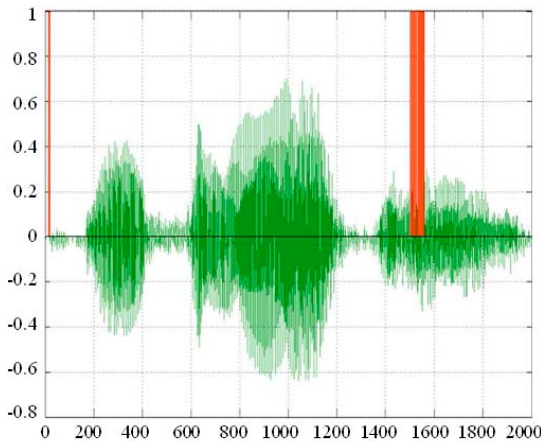


Fig. 7: Identified acute spot of aberration for the word Dinosaurs Using GA (Abnormal child)

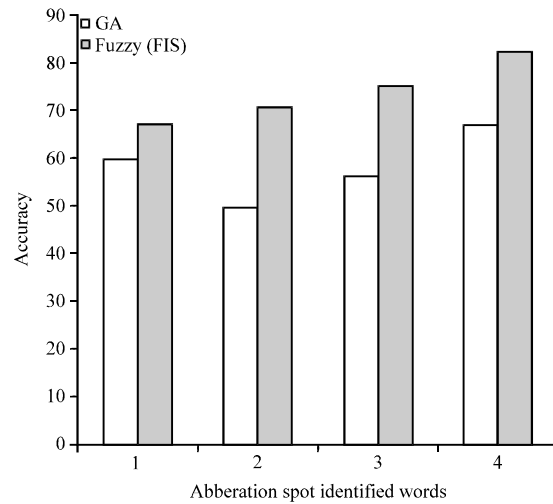


Fig. 9: Accuracy comparison for different abnormal words with existing Genetic Algorithm and Fuzzy Inference system

Finding spot using GA and FIS: In the previous work, it was tested with the database of 100 words with normal children two and an abnormal child data. Then the acute spot of aberration for the abnormal words is identified using GA and FIS separately. The output obtained for the words Dinosaurs using GA and FIS for identifying acute spot of aberration is shown in the Fig. 7 and 8, respectively.

Performance of GA and FIS for finding the aberration spot is calculated by TP, TN, FP and FN. Where TP, TN, FP and FN are:

True Positive (TP)-Aberration spot correctly identified and marked
True Negative (TN)-incorrect spot is identified as aberration spot
False Positive (FP)-Normal spot is identified as aberration spot
False Negative (FN)-Aberration spot is not identified as aberration spot

Performance measure for the GA and FIS for identifying the aberration spot is shown in Table 5 and

6, respectively. In this performance measure, as a sample four words were taken namely dinosaurs, flowers, daffodils and wild animals. Aberration spot was found for these words and the number of spots identified correctly and incorrectly are marked and accordingly TP, TN, FP and FN is calculated correspondingly. These values are shown in Table 5 and 6 of GA and FIS respectively.

Accuracy Comparison of GA and FIS for four different abnormal words is shown in Fig. 9 using the Table 5 and Table 6. The accuracy and MCC terms shows that performance of FIS is better than GA.

Table 5: Performance measure for the Genetic Algorithm for identifying the aberration spot

TP	TN	FP	FN	Sensitivity	FPR	Accuracy	Specificity	PPV	NPV	FDR	MCC
2	1	1	1	67	50	60	50	67	50	33	0.16
3	2	2	3	50	50	50	50	60	40	40	0.00
3	2	2	2	60	50	56	50	60	50	40	0.12
5	3	2	2	71	40	67	60	71	60	29	0.32

Table 6: Performance measure for the Fuzzy Inference System

TP	TN	FP	FN	Sensitivity	FPR	Accuracy	Specificity	PPV	NPV	FDR	MCC
2	2	1	1	67	33	67	67	67	67	33	0.33
4	1	1	1	80	50	71	50	80	50	20	0.32
2	1	1	0	100	50	75	50	67	100	33	0.57
6	3	1	1	86	25	82	75	86	75	14	0.63

CONCLUSION AND FUTURE WORK

The proposed system was implemented in the working platform of MATLAB (version 7.11). In this with the aid of the Free Audio Editor, the dataset with the normal and abnormal female children within the age limit 6-10 yrs are generated. For 100 normal data (words) 2 female children were utilized each and for 100 abnormal data, a female child was utilized for the system and their normal frequency range is from 20-4 kHz. The speech input is recorded at a sampling rate of 44.1 kHz. In this study, an effective system has been proposed to identify the abnormal word and the spot in which the speech has to be improved is also identified (Bharathi and Shanthi, 2011b, 2012d). The performance between ANN and SVM are compared and found that 90% accuracy is got by SVM and the performance of FIS is found better than GA.

REFERENCES

Bharathi, C.R. and V. Shanthi, 2011a. Classification of speech for clinical data using artificial neural network. *Int. J. Comput. Sci. Issues*, 8: 359-365.

Bharathi, C.R. and V. Shanthi, 2011b. Feature extraction using MFCC and survey on classification algorithms for clinical data. *Proceedings of the International Conference on Computer Science Engineering*, June 10-12, 2011, Shanghai, China.

Bharathi, C.R. and V. Shanthi, 2011c. Finding acute peaks and amplitudes of speech for clinical data using FFT. *Proceedings of the 1st International Conference on Advanced Computing Methodologies Conference*, December 9-10, 2011, Hyderabad andhra Pradesh, India.

Bharathi, C.R. and V. Shanthi, 2012a. An effective system for acute spotting aberration in the speech of abnormal children via artificial neural network and genetic algorithm. *Am. J. Applied Sci.*, 9: 1561-1570.

Bharathi, C.R. and V. Shanthi, 2012b. Discriminant analysis of disorder speech for clinical data. *Eur. J. Sci. Res.*, 90: 188-195.

Bharathi, C.R. and V. Shanthi, 2012c. Disorder speech classification for clinical data using SVM. *Proceedings of the National Conference on EXCITE 2012*, April 20-21, 2012, IETE Chennai Centre, Chennai, India.

Bharathi, C.R. and V. Shanthi, 2012d. Spotting the aberration spot in a speech with the aid of fuzzy inference system. *Indian J. Innovations Dev.*, 1: 795-802.

Bharathi, C.R. and V. Shanthi, 2012e. Survey on objective assessment of stuttered speech signal for disabled children. *Proceedings of the International Conference on Cloud Computing and eGovernance*, July 26-28, 2012, Bangkok, Thailand.