

An Overview of Subjective and Objective Assessment of Stuttering

Kusuma. H.R¹ and G. Seshikala²

¹Department of Electronics and communication, PES College of Engineering Mandya, India

Email: kusumaramanna.h@gmail.com

²School of ECE, REVA University Bengaluru, India

Email: seshikala.g@reva.edu.in

Abstract— Stuttering is a speech disorder, causes frequent problem with normal flow of speech. The normal fluency of the speech is disturbed by uncontrolled prolongation and repetitions of words, sounds, phrases or syllables and unwanted blockages in speech. Researchers are more interested in this field because in the field of speech language pathology and speech signal analysis, stuttering plays an important role. Stutter speech processing found its application in stuttering assessment and analysis that helps a speech language pathologist in the treatment and diagnosis. Thus, understanding of the stutter speech behavior by the researchers is a fundamental requirement and that helps to grasp the subject of interest and problem solving. Hence keeping this object in mind, this paper provides an insight to overview of the conventional assessment techniques and objective assessment techniques used for stuttering assessment. It's also provides insight to the further development in this area. The paper summarizes the different stuttering disfluency classification techniques and introduces the subjective and objective assessment of stuttering and its applications till current date. Summary of the classifiers table included at the end of the paper for better understanding. In addition, the previously used classification techniques are explained to make the paper self-sufficient and they are Artificial Neural Networks (ANN), Hidden Markov Model (HMM), k -nearest neighbor (k -NN), Linear Discriminate Analysis (LDA), and Support Vector Machine (SVM).

Index Terms— ANN, HMM, LPCC, MFCC, SVM, stuttering.

I. INTRODUCTION

Speech communication is an effective communication used by people to exchange the ideas and thoughts through a language. Normally the speech is produced with the repetition, hesitation, pauses and interjections, without these disruptions, speech doesn't sound good [2]. These disruptions in speech is normal, if they are not ruin the normal flow of speech like duration of speech and frequency of occurrence of the speech, if not considered it as abnormal speech. The abnormality in speech production is called as speech disorder or speech disfluency, out of many speech disfluencies, stuttering plays a most important role because stuttering causes severe problems on the quality of life of adults and children who stutter.

In stuttering the normal flow of speech is disturbed by abnormal repetitions (what-what happened), prolongation (www what happened), and abnormal stoppages or pauses (no sound) of sounds and syllables, it results disfluency

in speech. “All India Institute of Speech and Hearing” (AIISH) have been reported that 10% of the people suffering from the stuttering who having communication disorder. In worldwide more than 70 million people are affected by stuttering disorder which comes to 1% of the world population [5, 24, 39] Table I indicates different types of stuttering dysfluencies.

It is most common that children will stutter in the learning stage of speech and language processes is around age between 2 to 6 years. For some children it persists in the adulthood, because they are not able to get control over their speech production, they are underdeveloped phase and this type of stuttering is called as developmental stuttering. Majority of the brain injuries and brain diseases i.e. stroke, tumor etc. caused the stuttering and it is called as neurogenic stuttering. Psychological factors like, stress, emotional trauma causes effect on the thinking and reasoning this result disfluent speech and it is known as psychogenic stuttering.

TABLE I. TYPES OF STUTTER SPEECH DYSFLUENCIES

<i>SL. No</i>	<i>Types of Disfluency</i>	<i>Example</i>
1	Whole word repetition	Where where where are you going?
2	Part word repetition	Where r r r is it?
3	Phrase repetition	I need to I need to I need to do it
4	Sound/syllable Prolongation	Whooooo is it
5	Filled pauses	Umm-uhh-well. I am going to the store
6	Silent Pauses	I am cleaning the [pause] house
7	Broken words	I am clea[pause] ing to the house

In traditional stuttering treatment method, a Speech Language Pathologist (SLP) counts the dysfluencies, classify the abnormality and estimate the severity manually, to keep track further improvement in the quality of treatment. This type of stuttering treatment method is biased and subjective [5] and purely depends on the practice and experience of the Speech Language Pathologist (SLPs) may be causes the human error. Hence, to make it feasible and reliable SLPs need to facilitate with speech processing technology and artificial intelligence [8].

II. CLASSIFICATION OF SUBJECTIVE ASSESSMENT OF STUTTERING

A. Conventional Stuttering Assesment

Assessment plays very important role in stuttering treatment, in subjective assessment is done by the Speech Language Pathologist (SLP) or self-assessment by people who stutter (PWS). In which generally performance analysis is done before starting the treatment, during the treatment and following the treatment to find stuttering severity using rating scale [16]. In monitoring process, Speech Language Pathologist (SLP) will train the parent or guardian to assess the stuttering symptoms of children [17].

Speech Language Pathologist formally assesses speech using subjective and objective testing, the speech history includes a parental interview (Table II), during which attention is paid to parent-child interaction. To measure the severity of stuttering in adults and children’s “stuttering severity measuring instrument” is used and for young children the “stuttering prediction instrument” is used. These instruments provide the information about stuttering type, rate and duration and evaluate the entire speech to find whether any secondary characteristics are present to conclude the observation with proper therapy. Diagnostic instruments standardization is not based on the quality of findings, the quality of the findings are defines the level of certainty of an outcome measure and effect size using GRADE. GRADE is a method used to assign the score based on the quality of evidence and outcome. Effect of negative evaluation on speech therapy is depicted in the Table III.

TABLE II. QUESTIONS THAT WILL HELP PARENTS DISTINGUISH BETWEEN NORMAL SPEECH DYSFLUENCY AND STAMMERING

<i>Sl. No</i>	<i>Question</i>
1	Is the child repeating parts of words rather than whole words or phrases? (Example, “a-a-a-angel”)
2	Is the child repeating sounds more than once every eight to ten sentences?
3	Is the child repeating the same word more than two times?
4	When the child has trouble with a word, does he or she seem frustrated or embarrassed?
5	Does the child stutter for more than six months?
6	Do they use extra words or sounds like “uh” or “um” or “well” to begin words?
7	When the child tries to talk, does he ever get so stuck that he does not produce any sound for several second?
8	Are the child’s stuttering behaviors causing him or her to avoid talking to use substitute words.

B. Conventional Stuttering Validation

Validation is a subjective process done by the expert [17], internal consistency consider as a quality of individual items is measured by scale and assigning the “composite score”. “Test-retest reliability” estimates the steady score at two separate point of time; two available reliability estimation techniques are intra-rater and inter-rater. In “Intra rater reliability” the estimation depends on whether the evaluator assigns the same score to individual situations whether the test executes on various occasions [17] and in Inter-rater reliability estimates whether the two evaluator rate independently in the same way. Since intervention of human in the above mention assessment and validation techniques, there will be chances of biased evaluation, if it is negative, the risk factor is high. To overcome this problem, there will be a need of automatic assessment and validation technique [17].

TABLE III. EFFECT OF NEGATIVE EVALUATION ON STUTTER SPEECH THERAPY

Risk Factor	Grade	Consequence
Bias	1	Serious
	2	Very serious
Inconsistency	1	Serious
	2	Very serious
Indirect Evidence	1	Serious
	2	Very serious
Inaccuracy	1	Serious
	2	Very serious
Publication bias	1	Likely
	2	Very likely

Validation is a subjective process done by the expert [17], internal consistency consider as a quality of individual items is measured by scale and assigning the “composite score”. “Test-retest reliability” estimates the steady score at two separate point of time; two available reliability estimation techniques are intra-rater and inter-rater. In “Intra rater reliability” the estimation depends on whether the evaluator assigns the same score to individual situations whether the test executes on various occasions [17] and in Inter-rater reliability estimates whether the two evaluator rate independently in the same way. Since intervention of human in the above mention assessment and validation techniques, there will be chances of biased evaluation, if it is negative, the risk factor is high. To overcome this problem, there will be a need of automatic assessment and validation technique [17].

III. CLASSIFICATION OF OBJECTIVE ASSESSMENT OF STUTTERING

For a decade, speech processing in an area of signal processing, that has developed rapidly over the past decade. This rapid progress results significant advancement in automatic speech processing applications by using new machine learning paradigms [18]. The rapid growth in the automation technology in speech processing, there will be a major advancement in automatic stutter speech assessment in Speech Language Pathology. The research about automatic assessment started in 1995, Howell and Sackin classified major two types of stuttering events, prolongation and repetition using artificial neural network [14], with the hit or miss rate of 0.77 and 0.82 using autocorrelation function (ACF) and spectral coefficients, the main objectives are analysis and classification of repetition and prolongation in stuttered data base. This paper does not imply that single automatic dysfluency detection is the proper solution for all kind of stuttering problems. Indeed, the speaking rate and characteristics are differ in different language, while processing these signals at most attention need to be provided while extracting the features for classification.

A. Artificial Neural Network (ANN) Classifier

The ANN is a mathematical models used to develop a smart computing systems, it is inspired by the biological neural network. ANN is used to solve many problems in research areas like control system, signal prediction, pattern recognition of the signal and memory optimization [35]. ANN plays a major role in the speech processing [36] like speaker recognition [38] and it is an irreplaceable tool to find correlation between the similar signals [31, 21]. Due to its efficient performance ANN is a widely used in stutter type identification, such as whether it is a prolongation or repetition and also used to find the fluency and dysfluency percentage in the stuttered speech.

Artificial neural network has the exceptionable potential of information processing, due to its parallelism, fault and noise tolerance, learning and recognition capabilities. In the findings [19, 20] data of twelve children with stuttering (UCLASS database), and the extracted nine parameters are used to train the ANN. Manual

segmentation technique is used to extract the word, the ANN trained model successfully classified a word either repetition or prolongation. In [27] ANN rough set is used to detect the stutter events, vowel prolongation and syllable repetition.

The ANN model is also used in identification of stuttering events such as repetitions and prolongations. Most of the stutterer's suffer from repetition and prolongation problem so these two events play a major role in the stuttering [28] assessment. Stuttering type identification using ANNs [30, 20, 26] done by labeling with the labels repetition, prolongation and others type. The training process is done by providing the acoustic input features either repetitions or prolongations and output through a layer of hidden units. Figure 1 shows architectural flow graph of multilayer "feed forward neural network".

B. Hidden Markov Model (HMM) Classifier

Probabilistic model play a major role in real word statistical data analysis. HMM is a significantly used probabilistic model for analysis of prolongation and repetition in stutter speech due to speech signals are considered as piece wise linear. HMM is used to evaluate the degree of stuttering during the therapy session, for establishment of disfluency frequency, duration and speaking rate with a high correlation coefficient [22]. Stuttering assessment for non-English Malay language is used for analysis, the HMM can perform the statistical analysis like, counting and classification. Count the number of stuttering events and classifies event among the classes such as repetitions, word duration and pauses [23].

The statistical observation parameters, duration of disfluency, duration of the speech and rate of speech are used to find the stuttering severity. Noth [22] used samples of 37 patients who having a stuttering problem while reading an entire passage or beginning of a passage. Phoneme accuracy is calculated based on the number of detected dysfluencies results correlation coefficient of up to 0.99. However, model seems to over-estimate the reading errors and author concluded that findings are not sufficient and expect more research in the future, especially in the identification of stoppages and repetition.

Wiśniewski [32, 33] in 2007, presented 2 papers about "an automatic detection module using HMM". They used prolongation 38 samples and stoppages 30 samples to train the recognition model and the summary model trained with the 30 samples [38]. All the samples were decimated by the factor 22050 Hz and the parametric analysis is done by using MFCCs. With reference to their result, recognition accuracy of the summary models is 70%. An automatic detection system is developed to recognize the prolonged fricative sounds with HMM classifier [33]. Acoustic signals are parameterized before analysis, MFCC used as a characterization parameter with the sampling frequency of 22050 Hz and all the sound samples records were normalized to the same "dynamic range – 50dB", it showed the result of $\approx 80\%$.

Tian-Swee [39] proposed "automatic stuttering recognition system utilizes HMM technique" to evaluate stutter speech disorder. "Speech Therapy Assistance Tools (MSTAT)" helps speech language pathologist (SLP) to identify language and speech disorder in children. In HMM, Markov model used as a processing model and trained with the normal speech and disorder speech shape of the children. The statistical characteristics are probabilistically demonstrated through the definite observation [39].

In Markov Chain speech signals are considered as a "stochastic function" of the state sequence by considering state sequence is hidden. Following are the two types of HMM first "observation probability functions" model called as "discrete HMM (DHMM)" and second is "continuous density HMM (CDHMM)". As shown in the Figure 2 there are five states in the HMM "left-to- right model". 20 normal speech and 15 artificial stutter speech are considered in the reference data base. Each of 10 samples are used are used to train the speech model and each of 5 samples are used in test model of HMM. If the recorded value is greater than that of the threshold value, it is identified as a normal speech or else it is identified as artificial speech with the accuracy of 96% and 90%. In 2007, 3 other researchers are incorporate the same feature extraction and classification technique with different sampling frequency, states in model, training and testing data samples. Different criteria will results in a different accuracy on stuttering recognition.

C. Support Vector Machine (SVM) Classifier

SVM is powerful tool used for pattern recognition, in which separation of the classes is done using a hyper plane. They have used two major classes, fluent and disfluent using different decision logic [24, 40, and 41]. Binary classifier having optimal linear decision boundary using SVM used to distinguish the classes, in which the main design consideration is risk minimization. The decision making is on weighted sum of the elements in the training data set, these elements are considered as support vectors. These vectors are used separate the two different classes using a hyper plane [8].

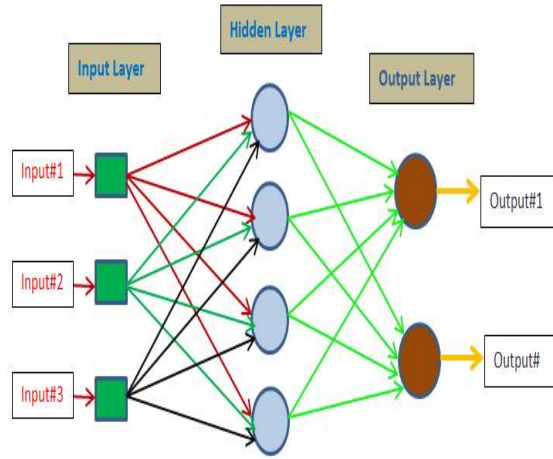


Figure 1. The basic architecture of multilayer feed forward neural network

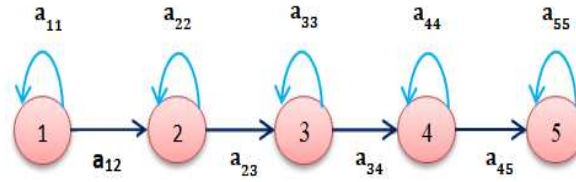


Figure 2. Representation of left-to-right HMM[39]

Ravikumar [24] proposed “automatic detection method for syllable repetition in reading speech for objective assessment for stuttered dysfluencies” design method consists of following four stages, segmentation, feature extraction, score matching and decision making. Decision logic is implemented using SVM classifier to classify fluent and/or dysfluent speech, for this they used data set consists of fifteen adult stutter speech samples, twelve samples used to train the model and three samples are for testing, which provides the accuracy of 94.35%. They got some comments on their works related to reliability of the results due to minimum data usage. To assure the reliability, system need to trained with the more number of training data, even though Ravikumar [13] got 83% of accuracy he only used 8 samples to train the system and 2 samples in testing so reliability of the outcome is questionable. In database they not considered the important features like origin, gender age[24, 13, 21], because these features plays an important role in measure of severity of stuttering. Tian-Swee [39] employed artificial stutter speech for testing and training the system, 10 samples for training and 5 samples for testing the system and got the accuracy of 90%, still the outcome was questionable due to they only employed the artificially generated speech to test the system they neglected the natural stutter speech and not mentioned how they generated the artificial speech. System not done the stutter speech classification even though they mentioned it was one of the objectives. The reliability of the outcome can be assured by “cross-validation and confidence intervals”, cross validation used to validate the classifier accuracy [42] by this performance of the system can be measured.

D. “*k*-Nearest Neighbour (*k*-NN)” Classifier

“*k*-NN” is a simplified classifier, which incorporated lazy learning [9]. It incorporated “supervised learning algorithm” for classification in which to find minimum distance, the distance between “query instance” (test speech signal) and each of the “training instances” (training speech signal) is calculated. The minimum distance is used to find the “*k*-NN” group. Each query instances is compared with all the training instances, the query instance prediction is based on nearest neighbor category voting, this takes high execution time [9, 10].

To predict each test speech signal, needs to find minimum distance between the “test speech signals” to each of the “training speech signal” in the training set to find the *k*-NN category of the training data set. Euclidean Distance is used to calculate the distance between each and every data in the training set with the test class to

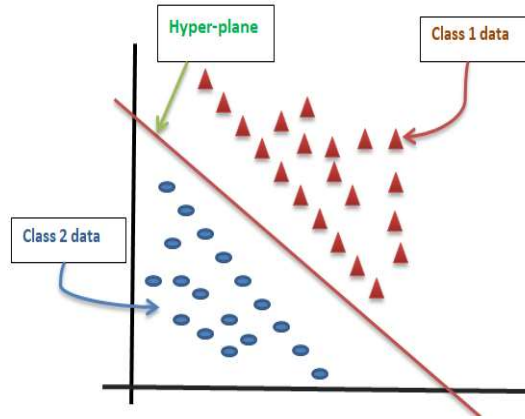


Figure 3. Decision making algorithm in SVM

examine how close each training set is to the test class. Euclidean Distance can be measured using the following mathematical model.

$$d_E(x, y) = \sqrt{\sum_{i=1}^n x_i^2 - y_i^2} \quad (1)$$

Classification of stutter events is done by assigning labels like repetitions or prolongations for every test event, by finding the k closest members.

E. Linear Discriminate Analysis (LDA) Classifier

Many signal processing applications like image retrieval, speech recognition, face recognition, etc. LDA used for data and feature extraction. Finding [9, 10] linear transformation is used to find the discrimination function to classify two feature vector classes. Linear transformation is a transformation technique used to transform each “speech feature vector” into a “projection feature vector”. The following mathematical model is used to find the discriminant function ‘ f_i ’

$$f_i = \mu_i S_w^{-1} x_k^T - \frac{1}{2} \mu_i S_w^{-1} \mu_i^T + I_n(P_i) \quad (2)$$

$\mu_i \rightarrow$ Mean of i^{th} group feature vector; where $i=1, 2$

$X_k \rightarrow k^{\text{th}}$ feature representation of X .

$P_i \rightarrow$ Sum of sample in each group is divided by the total number of samples.

Mean of the entire data set μ_3 is defined by the equation (3), where group 1 mean represented by μ_1 , group 2 mean represented by μ_2 . P_1, P_2 are probability estimation of the groups 1 and 2 respectively.

$$\mu_3 = P_1 \times \mu_1 + P_2 \times \mu_2 \quad (3)$$

In LDA, class discrimination is done within the group S_w (Scatter within-group) and between groups S_b (Scatter between group). Scatter within group is defined by

$$S_w = P_1 \text{Cov}_1 + P_2 \text{Cov}_2 \quad (4)$$

Covariance matrix is defined by the equation (5), where Cov_1 and Cov_2 are the estimated covariance's of groups 1 and 2. Since Cov_1 and Cov_2 are symmetric, equation (6) is used to find the scatter between-group. S_w is considered as a covariance of data set by replacing “mean vectors” of each group with “mean vectors” of total data set.

$$\text{Cov}_i = (x_i - \mu_i)(x_i - \mu_i)^T \quad (5)$$

$$S_b = \sum_i (\mu_i - \mu_3) \times (\mu_i - \mu_3)^T \quad (6)$$

Transform the extracted data features into “discriminant function”, training data set and the prediction data sets are mapped into new coordinate. The difference between ‘I’ and ‘f’ is considered as class label prediction data [9, 10].

IV. APPLICATIONS OF STUTTER SPEECH MONITORING

It is used in speaker monitoring [1]; it helps stutter patient for self-evaluation. It helps Speech language pathologist during treatment [5]. It helps in avoid the judgment biasing during the assessment by expert or a guardian. Machine learning and deep learning incorporation in stuttering assessment improves the accuracy of the automatic detection and classification. It is used to measure the degree of speech disability [6]. Removal of dysfluency from the stutter speech helps anti stuttering applications like automatic speech recognition (ASR). Used in, “delayed auditory feedback” (DAF) and “Frequency auditory feedback” (FAF) systems [7]. It helps to distinguish different stutter dysfluencies [9]. It is very difficult to speech language pathologist (SLP), to keep track of the treatment manually. The manual assessment techniques are inconsistency and time consuming and most of the time gives erroneous outcome [9]. To overcome this problem there is a need of automatic assessment and recognition technique to improve the efficiency.

V. CONCLUSION

Stutter speech assessment demands efficient system to automate the assessment process for analysis and classification. This paper mainly focuses on subjective assessment of stuttering, it gives strong understanding of conventional stuttering assessment and validation methods. This gives insight to the objective assessment of stuttering, with a note on the reviewed papers currently existing techniques for stutter speech classification, feature extraction and applications are viewed. By keeping this as a base study, advancement in this field can be studied. Further development in stuttering assessment, machine learning, deep learning and fusion of feature extraction techniques are scope for research. The key success rate between subjective assessment and objective assessment using machine learning based fusion feature extraction technique yet to explore. Computational complexity is also an area for analysis. Table 4 shows the summary of the review literature.

TABLE IV. SUMMARY OF THE REVISED LITERATURE ON AUTOMATIC STUTTER SPEECH RECOGNITION SYSTEM

<i>Author</i>	<i>Year</i>	<i>Database</i>	<i>Features</i>	<i>Classifier</i>	<i>Result in %</i>
Howell	1995	-	Envelope parameters and Autocorrelation function	ANN	80%
Howell	1997	12 samples	Energy, Maxima's, Duration, Spectra of word and Spectra of part word	ANN	78.01%
Geetha	2000	51 samples	Age, Sex, Type of dysfluency, Frequency of dysfluency, Rate of speech, Duration of speech, Behavioral scores, Physical concomitant, Family history, Historical and Attitudinal	ANN	92%
Nöth	2000	37 samples	Dysfluen duration, frequency and rate of speech	HMM	-
Czyzewski	2003	6 samples of normal and 6 samples of stop-gap speech	Frequency of 1 st and 3 rd formant's Frequency and Amplitude	ANN and rough set	73.25% $\geq$ 90.0%
Szczurowska	2006	8 samples	512-point FFT	MLP Kohonen	76.67%
Wiśniewski	2007	38 samples of prolongation 30 samples of stop blockade and 30 samples of silence	MFCC	HMM	70%
Wiśniewski	2007	-	MFCC	HMM	Approximately 80%
Tian-Swee	2007	15 samples of normal speakers, 10 samples of artificial stuttered speech	MFCC	HMM	96%
Ravikumar	2008	10 samples	MFCC	Perceptron	83%

Ravikumar	2009	15 samples	MFCC	SVM	94.35%
Świetlicka	2009	8 samples of stuttering speakers, 4 samples of normal speakers(59 fluent speech and 59 non-fluent speech)	512-Point FFT	Kohonen MLP RBF	88.1% - 94.9%
Lim Sin Chee	2009	92 male recording, 15 female recording, 38 male speakers, 5 female speakers	Linear Prediction Cepstral Coefficient(LPCC)	LDA k-NN	87.50% 89.77%
Lim Sin Chee	2009	92 male recording, 15 female recording, 38 male speakers, 5 female speakers	Mel Frequency Cepstral Coefficients (MFCC)	LDA k-NN	90%
P. Mahesha	2015	100 repetition, 100 prolongation dysfluency segments	MFCC, DMFCC, DDMFCC, Pitch, Energy, and Duration	SVM	96.85%
P. Mahesha	2017	500 speech samples are used for training and testing	MFCC and LH-MFCC	GMM	93.19%
G. Manjula	2017	15 to 35 years age group (16 male, 4 female speakers)	Epoch	-	-
M Manjutha	2020	2,926 speech samples	PSO, SFO, MFCC	SVM NB	88.60% 95.21% - 97.01%

ACKNOWLEDGEMENT

“We acknowledge the support extended by VGST, Govt. of Karnataka, in sponsoring this research work with Software tools and database vide ref. No. KSTePS/VGST/GRD-684/KFIST(L1)/2018”.

REFERENCES

- [1] Tedd Kourkounakis, Amirhossein Hajavi, Ali Etemad, “Detecting speech disfluencies using a deep residual network with Bidirectional Long short-term memory,” ICASSP 2020, 978-1-5090-6631-5/20/\$31.00 ©2020 IEEE.
- [2] Fathima Afroz, Shashidhar G Koolagudi “Recognition and classification of pauses in Stuttered speech using Acoustic Features,” 2019 6th International Conference on Signal Processing and Integrated Networks (SPIN) 978-1-7281-1380-7/19/\$31.00 ©2019 IEEE.
- [3] M Manjunath, P Subashini, M Krishnaveni, V Namadha, “An Optimal Cepstral Feature Selection Method for Dysfluencies Classification using Tamil Speech Dataset”, 5th IEEE International Smart Cities Conference (ISC2 2019).
- [4] Ankit Dash, Nikhil Subramani, Tejas Manjunath, Vishruti Yaragarala and Shika Tripathi, “Speech Recognition and Correction of a Stuttered Speech,” 978-1-5386-5314-2/18/\$31.00 ©2018 IEEE.
- [5] P. Mahesha, D. S. Vinod, “LP-Hilbert Transform Based MFCC for Effective Discrimination of Stuttering Dysfluencies,” This full-text paper was peer-reviewed and accepted to be presented at the IEEE WiSPNET 2017 conference.
- [6] Gunjan Jhawar, Prajacta Nagraj, P. Mahalakshmi, “Speech Disorder Recognition using MFCC,” International Conference on Communication and Signal Processing, April 6-8 2016, India.
- [7] Antonio C. Bortoletto, Mario Minami, Celso S. Kurashima, “DSP Altered Feedback System for Anti- Stuttering Applications,” 2016 IEEE International Symposium on Consumer Electronics.
- [8] P. Mahesha, D. S. Vinoda, “Combining Cepstral and Prosody Features for Classification of Disfluencies in Stutter Speech,” 1 Springer India 2015 L.C. Jain et al. (eds.), Intelligent Computing, Communication and Devices, Advances in Intelligent Systems and Computing 308, DOI 10.1007/978-81-322-2012-1_67.

- [9] Lim Sin Chee, Ooi Chia Ai, M. Hariharan and Sazali Yaacob, " MFCC based Recognition of Repetition and Prolongation in Stuttered Speech using k-NN and LDA," Proceedings of 2009 IEEE Student Conference on Research and Development (SCORED 2009), 16-18 Nov. 2009, UPM Serdang Malaysia.
- [10] L. S. Chee, O. C. Ai, M. Hariharan, and S. Yaacob, "Automatic detection of prolongations and repetitions using LPCC," 2009, pp. 1-4.
- [11] Vicky Zayats, Trang Tran, Richard Wright, Courtney Mansfield Mari Ostendorf, "Disfluencies and Human Speech Transcription Error," arXiv: 1904.04398v1 [cs. CL] 8 Apr 2019.
- [12] Elzbieta Szabelska, Anna Kruczynska, "Computer- Based Speech Analysis in Stutter," speech therapy, stutter, logopaedics, signal analysis.
- [13] K. M. Ravikumar, Balakrishna Reddy, R. Rajagopal, and H. C. Nagaraj, "Automatic Detection of Syllable Repetition in Read Speech for Objective Assessment of Stuttered Disfluencies," World Academy of Science, Engineering and Technology International Journal of Electrical and Computer Engineering Vol:2, No:10, 2008.
- [14] M.Hariharan, Vikneswaran Vijejan, C.Y. Fook, Sazali Yaacob, "Speech Stuttering Assessment Using Sample Entropy and Least Square Support Vector Machine," 2012 IEEE 8th International Colloquium on Signal Processing and its Applications.
- [15] G. Manjula, M. Shiva Kumar, Y. V. Geetha, T. Kasar, "Identification and Validation of Repetitions/Prolongations in Stuttering Speech using Epoch Features," International Journal of Applied Engineering Research ISSN 0973-4562 Volume 12, Number 22 (2017) pp. 11976-11980 © Research India Publications. <http://www.ripublication.com>.
- [16] Jeanna Riley, Glyndon Riley, Gerald Maguire, "Subjective Screening of Stuttering severity, locus of control and avoidance: research edition," Journal of Fluency Disorders 29 (2004) 51–62.
- [17] Pertjys, M.A.J., Oonk, L.C., Beer, de J.J.A., Bunschoten, E.M., Bast, E.J.E.G., Ormondt, van J., Rosenbrand, C.J.G.M., Bezemer, M., Wijngaarden, van L.J., Kalter, E.J., Veenendaal, van H. "Clinical Guideline Stuttering in Children, Adolescents and Adults". NVLF, Woerden, (2014).
- [18] Louis Przebienda, M.S. Richard Merson, "The Application of Digital Speech Processing to Stuttering Therapy," IEEE Instrumentation and Measurement Technology Conference Ottawa, Canada, May 1991, 1997.
- [19] P. Howell, S. Sackin, and K. Glenn, "Development of a Two-Stage Procedure for the Automatic Recognition of Dysfluencies in the Speech of Children Who Stutter: I. Psychometric Procedures Appropriate for Selection of Training Material for Lexical Dysfluency Classifiers," J Speech Lang Hear Res, vol. 40, pp. 1073-1084, October 1, 1997.
- [20] P. Howell, J. Au-Yeung, S. Sackin, K. Glenn, and L. Rustin, "Detection of supralexical dysfluencies in a text read by children who stutter," Journal of fluency disorders, vol. 22, pp. 299-307, 1997.
- [21] A. Czyzewski, A. Kaczmarek, and B. Kostek, "Intelligent processing of stuttered speech," Journal of Intelligent Information Systems, vol. 21, pp. 143-171, 2003.
- [22] E. Nöth, H. Niemann, T. Haderlein, M. Decher, U. Eysholdt, F. Rosanowski, and T. Wittenberg, "Automatic stuttering recognition using hidden Markov models," 2000.
- [23] T. S. Tan, A. Ariff, C. M. Ting, and S. H. Salleh, "Application of Malay speech technology in Malay speech therapy assistance tools," 2007, pp. 330-334.
- [24] K. Ravikumar, R. Rajagopal, and H. Nagaraj, "An approach for objective assessment of stuttered speech using MFCC features," ICGST International Journal on Digital Signal Processing, DSP, vol. 9, pp. 1924, 2009.
- [25] Lim Sin Chee, Ooi Chia Ai, Sazali Yaacob, "Overview of Automatic Stuttering Recognition System", Proceedings of the International Conference on Man-Machine Systems (ICoMMS) 11 – 13 October 2009, Batu Ferringhi, Penang, MALAYSIA.
- [26] J. MacClelland and D. Rumelhart, "Parallel distributed processing: Explorations in the microstructure of cognition": The MIT Press, 1988.
- [27] I. Szczurowska, W. Kuniszyk-Józkowiak, and E. Smolka, "The application of Kohonen and Multilayer Perceptron Networks in the speech nonfluency analysis," Archives of Acoustics, vol. 31, p. 205, 2006.
- [28] M. Wingate, "The structure of stuttering: A psycholinguistic analysis": Springer Verlag, 1988.
- [29] P. Howell and S. Sackin, "Automatic recognition of repetitions and prolongations in stuttered speech," in Proceedings of the First World Congress on Fluency Disorders, 1995, pp. 372-374.
- [30] P. Howell and S. Sackin, "Automatic recognition of repetitions and prolongations in stuttered speech," in Proceedings of the First World Congress on Fluency Disorders, 1995, pp. 372-374.
- [31] Y. V. Geetha, K. Pratibha, R. Ashok, and S. K. Ravindra, "Classification of childhood disfluencies using neural networks," Journal of fluency disorders, vol. 25, pp. 99-117, 2000.
- [32] M. Wiśniewski, W. Kuniszyk-Józkowiak, E. Smolka, and W. Suszyński, "Automatic Detection of Disorders in a Continuous Speech with the Hidden Markov Models Approach," in Computer Recognition Systems 2. vol. 45/2008: Springer Berlin / Heidelberg, October 18, 2007, pp. 445-453.
- [33] M. Wiśniewski, W. Kuniszyk-Józkowiak, E. Smolka, and W. Suszyński, "Automatic Detection of Prolonged Fricative Phonemes with the Hidden Markov Models Approach," Journal of Medical Informatics & Technologies, vol. 11/2007, 2007.
- [34] I. Świątlicka, W. Kuniszyk-Józkowiak, and E. Smolka, "Artificial Neural Networks in the Disabled Speech Analysis," in Computer Recognition System 3. vol. 57/2009: Springer Berlin / Heidelberg, May 12, 2009, pp. 347-354.
- [35] A. K. Jain, M. Jianchang, and K. M. Mohiuddin, "Artificial neural networks: a tutorial," Computer, vol. 29, pp. 31-44, 1996.

- [36] W. Chen, S. Chen, and C. Lin, "A speech recognition method based on the sequential multi-layer perceptrons," *Neural Networks*, vol. 9, pp. 655-669, 1996.
- [37] K. Farrell, R. Mammone, and K. Assaleh, "Speaker recognition using neural networks and conventional classifiers," *IEEE Transactions on speech and audio processing*, vol. 2, pp. 194-205, 1994.
- [38] L. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proceedings of the IEEE*, vol. 77, pp. 257-286, 1989.
- [39] T. Tian-Swee, L. Helbin, A. K. Ariff, T. Chee-Ming, and S. H. Salleh, "Application of Malay speech technology in Malay Speech Therapy Assistance Tools," in *Intelligent and Advanced Systems, 2007. ICIAS 2007. International Conference on*, 25-28 Nov 2007, pp. 330-334.
- [40] R. A. El-Khoribi, "Support Vector Machine Training of HMT Models for Land Cover Image Classification}," *ICGST International Journal on Graphics, Vision and Image Processing, GVIP*}, vol. 8, pp. 7-11, December 2008.
- [41] C. J. C. Burges, "A Tutorial on Support Vector Machines for Pattern Recognition," *Data mining and knowledge discovery*, vol. 2, pp. 121-167, 1998.
- [42] M. Mullin and R. Sukthankar, "Complete cross-validation for nearest neighbor classifiers," in *Proceedings of the Seventeenth International Conference on Machine Learning*, 2000, pp. 639 – 646.