Gender Voice Recognition Using Random Forest Recursive Feature Elimination with Gradient Boosting Machines

Kudakwashe Zvarevashe
ICT and Society Research Group,
Durban University of Technology,
P.O. Box 1334, Durban 4000, South Africa
kudakwashe.zvarevashe@gmail.com

Oludayo O. Olugbara
ICT and Society Research Group,
Durban University of Technology,
P.O. Box 1334, Durban 4000, South Africa
oludayoo@dut.ac.za

***Abstract*—Speech emotion recognition is a difficult task in the field of affective computing because emotions in speech heavily depend on a variety of factors such as feeling, thought, behaviour, mood, temperament, personality and disposition that are hard to model. Emotion plays a significant role in decision making and it influences human perception, learning, behaviour and relationships between individuals. Gender voice is a contributing factor in boosting the accuracy of emotion recognition systems using speech signals. In this paper, we propose a gender voice recognition method which makes use of feature selection through the Random Forest Recursive Feature Elimination (RF-RFE) algorithm with Gradient Boosting Machines (GBMs) algorithm for gender classification. The training and testing data were obtained from a public gender voice dataset. The GBMs algorithm was later evaluated against the feed forward neural network and extreme machine learning algorithms. The classification accuracy of the GBMs improved after applying the RF-RFE to the dataset. Experimental results indicate that GBMs outperformed all the comparative algorithms in classification accuracy and proved to be a suitable candidate for gender voice recognition.**

***Keywords*— Affective computing; Gender voice recognition; Gradient boosting machines; Recursive feature elimination algorithm; Speech emotion recognition**

I. INTRODUCTION

Speech is one of the most popular ways that human beings use to establish communication [1]. It is used to express emotions together with a host of expressions that people use in daily activities. Expressing emotions is one of the primary features in human social interactions across different cultures [2]. Some researchers have argued that emotional cues stimulate human interactions and do carry more information when compared to the actual spoken words [3]. Speech processing has over the years given rise to research in voice biometrics [4], emotion recognition, automatic identification of age and gender [5] to boost customer satisfaction and relationships among other important reasons.

In addition, gender recognition plays a crucial role in Automatic Speaker Recognition (ASR) systems because gender-dependent systems perform better than gender-independent systems [6, 7, 8]. Gender recognition is a method that is frequently used to determine the gender category of a speaker through the analysis of speech signals [9]. This technique presents a plethora of opportunities for application areas such as speech emotion recognition, automatic speaker recognition, human to machine interaction and sorting of telephone calls by gender categorisation in the case of gender sensitive surveys [10]. Moreover, gender recognition systems can be used to automatically transfer phone call of an individual to the appropriate person or section [11]. Mobile healthcare systems can also benefit from this technique, especially in the case of a massaging service that requires a client to be massaged by the right person within the business environment [12].

Emotion expressions are based on the acoustic characteristics of a speaker that are highly influenced by gender and language of the speaker [13]. In [14], it was noted that application of gender recognition in Speech Emotion Recognition (SER) systems significantly boosts the performance of the systems. However, the development of SER systems that are independent of gender categorisation presents enormous challenges [15]. This is because emotional states are related to a couple of factors that include mentality and personality, which are quite difficult to model [7]. Pitch is one of the common features used in emotion recognition, but it varies over time because it is influenced by the emotional state of a speaker and is inspired by the need of an individual to adapt to the context of the subject driving the conversation [16].

Many researches have been done in gender recognition for over two decades, but it is still considered to be a challenging problem. Most of the challenges emanate from the fact that gender information is time-varying, independent of phonemes coupled with the issues of identity and independence for speakers who fall within the same gender category [17]. In [18], it was stated that gender recognition cannot reach an accuracy of 100% and it was proven in [19] that discriminating between male and female speakers is much more difficult than separating speech utterances and music. It was conveyed in [9] that most gender recognition systems are tested in clean speech.

The essential contributions of this paper lie in the feature ranking and selection methods applied as well as the

978-1-5386-3060-0/18/$31.00

recognition algorithm used. The proposed method for gender recognition in this paper is further evaluated against the state-of-the-art methods to validate its performance. The related literature is discussed in Section II. Section III describes the public database used for experimentation in this study while Section IV discusses the study methodology. In Section V, the results of the experiments conducted are presented and the paper is succinctly concluded in Section VI.

II. RELATED WORK

A variety of methods have been explored for gender recognition through the use of human physiological features such as face [20], length of the vocal folds [21] and features that have a substantive footprint in behaviour like speech [21], gait [22, 23] and lip movement [24] amongst others. In [21], pitch was used to detect gender by analysing the non-stationary behaviour of speech signals. They developed a peak detection algorithm to determine the peak peach and used the standard Fast Fourier Transform (FFT) technique to detect gender. They found that the recognition accuracy decreases faster in a frame based algorithm and that non-stationary component of speech signal increases gender detection accuracy. They used the Percentage Of Correctness (POC) parameter to evaluate the accuracy of the proposed algorithm was found to perform better than the state-of-the-art algorithms.

Maka et al. [25], discovered that non-linear smoothing increases the classification accuracy by 2%. They used the Texas Instruments Massachusetts Institute of Technology (TIMIT) corpus with 630 speakers, 438 males and 192 females. The experiment was conducted using different acoustic environments. They used the following features in the speech parameterisation stage. Mel-Frequency Cepstral Coefficients (MFCCs, 12 trajectories), Linear Frequency Cepstral Coefficients (LFCC, 12 trajectories), Linear Prediction Coefficients (LPC, 12 trajectories) and fundamental frequency estimated with YIN pitch detector (F0, one trajectory), Formants position (FMNT, 3 trajectories) and Formants bandwidth (FMNTBW, 3 trajectories). The correlation-based feature subset selection with the best first search strategy was used as a feature selection tool to reduce dimensionality. The recognition obtained was 99.4%, but they were silent on the recognition algorithm that was used.

An efficient time domain based gender detection algorithm that employs the autocorrelation as a pitch detection technique with the K-means algorithm was proposed in [26]. The authors created their own corpus consisting people who fall in the age group of 20-30 to conduct the experiment. The training consisted of 20 females and 20 males. They discovered that when the speech characteristics move from stationary to non-stationary, that is when the frame length increases, the detection rate decreases. Their study revealed that the hamming window has maximum accuracy of gender recognition and the K-means algorithm showed a higher accuracy.

In Archana and Malleswari [27], 5 males and 5 females were used to create an experimental corpus. In total, 15 utterances were collected from each speaker using the MATLAB software. The extracted features from audio signals were MFCCs, energy, entropy, median and standard deviation that include both time and frequency domain features. The Support Vector Machine (SVM) [1, 14, 28,] and Artificial Neural Network (ANN) machine learning methods were tested for classification. The final results show that SVM dominated ANN in gender classification of speakers.

Bocklet et al. [29] fused the vectors for age and gender recognition to come up with super vectors using the Gaussian Mixture Models (GMM) algorithm [30]. They used the MFCCs extracted from the German SpeechDat II corpus, which is annotated with gender and age labels as given by callers at the time of recording. They also used the VoiceClass corpus for further experiments. A Universal Background Model (UBM) was created using the Expectation Maximization (EM) algorithm. The created super vectors were later used with 3 different SVM kernels [1, 14, 28,], which are polynomial kernel, Radial Basis Function (RBF) kernel and linear GMM distance kernel, based on the Kullback Leibler divergence. The SVM [1, 14, 28,] improved the recognition rate to 74% ($p < 0.001$), which is considered to be in the same range as human beings. Their results produced an accuracy rate, which was 43% more than the Parallel Phoneme Recognizer (PPR), which is a system that was developed for ASR and automatic acoustic Language IDentification (LID). It was used for comparison purposes in age and gender recognition for telephone applications.

The method proposed by Kotti and Kotropoulos [7] outperformed the previous state-of-the-art methods, which used the best-first search technique through the feature space using the Naive Bayes classifier, which improves accuracy by 9.74%. They used two emotional databases that are the Berlin database of emotional speech and Danish emotional speech database. They selected 15 features from 1379 extracted features using the branch and bound feature selection algorithm. Ramdinmawii and Mittal [31], selected pitch using autocorrelation, signal energy and MFCCs as the base features for gender recognition. The TIMIT database was used in the experiment with the linear SVM as a classifier. Their results showed that MFCC outperformed the pitch and signal energy. MFCC had an accuracy of 69.23%, while accuracies of pitch and signal energy were below 60%.

In [32], it was demonstrated that each speech signal carries gender discriminatory information using the GMM [30] together with MFCCs. Alhussein et al. [11], proposed a feature that determines voice intensity using the Modified Voice Contour (MVC), time domain and provides a single value in the form of the area under the MVC. The calculation of the area under MVC was done using the Simpson's rule. They used 2 databases, which are the TIMIT database and the Arabic database. Their model showed great promise and they even compared SVM [1, 14, 28,] with GMM and the former proved to be better than the latter.

Zeng et al. [9], proposed the use of GMM [30] using a combination of parameters, which are pitch and 10th order relative spectral perceptual linear predictive coefficients to distinguish between male and female speeches. The classification accuracy was well above 90% for both noisy and clean speeches. In [33], it was discovered that MFCCs feature extraction is not the best technique to use in noisy environments. They proposed an algorithm that involves MFCCs with SVMs for gender recognition. They used two variations of SVM kernels [1], the RBF and polynomial. The RBF kernel outperformed the polynomial kernel. A two level GMM algorithm was proposed in [34] to detect age and gender. The detection accuracy was well above 90%. Jayasankar et al. [35], used the Genetic Algorithm (GA) to create an optimal feature subset, which improved accuracy significantly.

III. CORPUS

Acoustic features were collected from 1584 males and 1584 females. These features were then used to create a database that was modelled into a csv file in [36]. The database consists of 22 acoustic features as shown in Table I, which were used in this study.

TABLE I. GENDER VOICE DATASET FEATURES

ACOUSTIC FEATURES	
PROPERTY	**DESCRIPTION**
duration	length of signal
meanfreq	mean frequency (in kHz)
sd	standard deviation of frequency
median	median frequency (in kHz)
Q25	first quantile (in kHz)
Q75	third quantile (in kHz)
IQR	interquantile range (in kHz)
skew	skewness
kurt	kurtosis
sp.ent	spectral entropy
sfm	spectral flatness
mode	mode frequency
centroid	frequency centroid
peakf	peak frequency
meanfun	average of fundamental frequency measured across acoustic signal
minfun	minimum fundamental frequency measured across acoustic signal
maxfun	maximum fundamental frequency measured across acoustic signal
meandom	average of dominant frequency measured across acoustic signal
mindom	minimum of dominant frequency measured across acoustic signal
maxdom	maximum of dominant frequency measured across acoustic signal
dfrange	range of dominant frequency measured across acoustic signal
modindx	modulation index

IV. METHODOLOGY

This section presents the process for the proposed gender voice recognition system. The Recursive Feature Elimination (RFE) algorithm based on Random Forest (RF) was used as a dimensionality reduction measure. The RFE technique, as applied in the caret package, implements a backward selection of the acoustic features by ranking their importance to an initial model using all the predictors [37]. The RFE selection method [38] is essentially a recursive process that grades features according to a certain degree of importance. It is a greedy optimization procedure used to find the superlative performing subset of features.

The method of RF-RFE performs well, according to experiments done in [37] and it was concluded that intrinsic properties of features concerning relevance, redundancy and complementarity can be catered for using weights fusion. Fig 1 shows the overall process of gender voice recognition as proposed in this study.

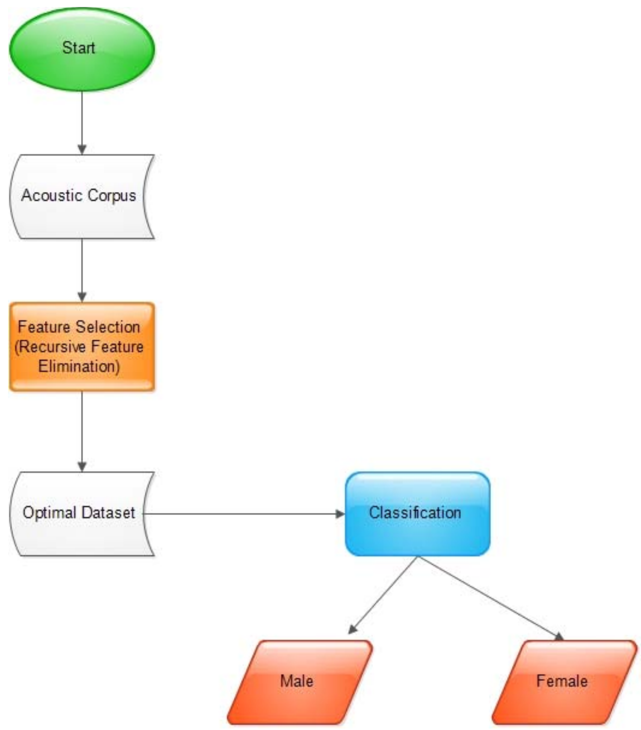

Fig. 1. The proposed gender voice recognition process of RF-RFE

Inputs:

Training set Tr

Set of α features Fe= $\{f_1........f_\alpha\}$

Ranking method M (Tr, Fe)

Outputs:

Final ranking R

Code:

Repeat for i in {1: α}

Rank set Fe using M (Tr, Fe)

$f^* \leftarrow$ last ranked feature in Fe

R $(\alpha - i + 1) \leftarrow f^*$

Fe$\leftarrow$ Fe $- f^*$

Fig. 2. Summarised RF-RFE algorithm

According to Fig 2, feature importance is measured at each iterative cycle and less relevant features are removed in each cycle. There is a greater deal of recursion in the process because for some measures, the relative importance of each feature can change considerably when assess on different subsets of features during the iterative purging process, especially in the case of highly correlated features. The final ranking is constructed following the order in which features are eliminated. The feature ranking method that we use in the process was given by RF. After creating a feature subset with the RF-RFE feature selection algorithm, the Feed Forward Neural Network (FFNN) classifier was applied.

A. *Feed Forward Neural Networks*

The FFNN is a branch of ANN, which is a mathematical model that applies the functions of the human brain to solve complex problem [39]. The inputs to the ANN are first multiplied by some link weights and are joined through a summation function. These are then handed over to a transfer function to create output neurons. The transfer function is applied to the weighted sum of the input neurons and the sigmoid function is the frequently used transfer function [40]. A FFNN is a biologically inspired classification algorithm, which comprises of a group of neuron like processing elements that are structured in a layered architecture. The connected neurons in the networked layer have different weights that encode the knowledge of the network.

B. *Gradient Boosting Machines*

The Gradient Boosting Machines (GBMs) algorithm was also used as a classifier. The GBMs are a family of potent machine learning methods that have performed successfully in an extensive assortment of practical applications [41]. In the GBMs algorithm, both the loss function and base learner models can be specified arbitrarily. The solution to the required parameter estimates can be very difficult to obtain given some specific loss functions $\Psi(y,f)$ and $h(x,\theta)$. A new function was proposed to solve this problem. The function $h(x,\theta_t)$ was proposed to be the most parallel to the negative gradient $\{get(x_i)\}^{N}_{i=1}$ } along the observed data:

$$g_t(x)=\left[\frac{\partial\psi(y,f(x))}{\partial f(x)}\right]_{f(x)=f_{t-1}(x)} \quad (1)$$

A function can be chosen in order to improve the correlation with gt(x). This creates room for the removal of a potentially extremely hard optimization job with the classical least squares minimization:

$$(\rho_t,\theta_t)= \arg\min_{\rho,\theta}\sum_{i=0}^{n}\left[-g_t(x)+\rho h(x_i,\theta)\right]^2 \quad (2)$$

As proposed by Friedman [42], the algorithm can be summarised as shown in Fig 3. The precise state of the algorithm, including all the necessary formulae will deeply rely on the design choices of $\Psi(y,f)$ and $h(x,\theta)$.

GRADIENT BOOSTING MACHINES ALGORITHM

Inputs:
Data $(y,f)^{N}_{i=1}$
Number of iterations M
Choice of the loss function $\Psi(y,f)$
Choice of the base learner model $h(x,\theta)$
The Algorithm:
Initialize f_0 with a constant
For t = 1 to M do
  Compute the negative gradient $g_t(x)$
  Fit a new base-learner function $h(x,\theta_t)$
  Find the best gradient descent step-size ρ_t:
  $\rho_t = \arg\min_{\rho} \sum_{k=0}^{n} \Psi[y_i, f_{t-1}(x_i) + \rho h(x_i,\theta_t)]$
  Update the function estimate:
  $f_t \leftarrow f_{t-1} + \rho_t h(x,\theta_t)$
End for

Fig. 3. Summarised GBMs algorithm

V. RESULTS AND DISCUSSION

The FFNN classifier, GBMs algorithm and Extreme Learning Machines (ELMs) have great precision when it comes to binary classifications. The first experiment was done without applying feature selection. All the classifiers achieved accuracies that were above 90%. This shows great promise when compared to the classic algorithms such as the Hidden Markov Model (HMM) [34] and GMMs [14] used by the previous researchers. As shown in Table II and Fig 4, the GBMs algorithm had the highest accuracy (97.58) followed by FFNN (96.95) with the ELMs algorithm having the least accuracy value of 96.53. The GBMs algorithm also had the highest Kappa valued, followed by FFNN then ELMs.

TABLE II. RESULT BEFORE APPLYING RF-RFE

Method	Accuracy (%)	Kappa (%)
FFNN	96.95	93.89
ELMs	96.53	93.05
GBMs	97.58	95.16

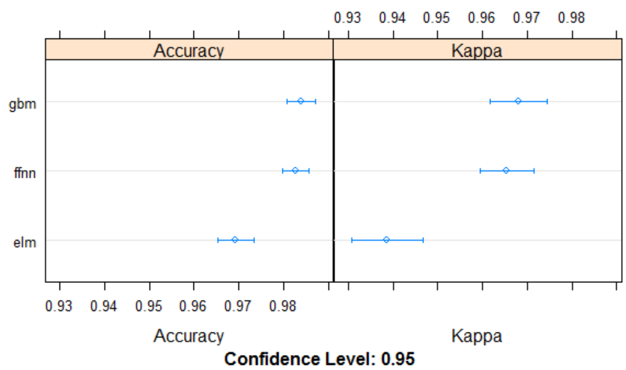

Fig. 4. Accuracy and Kappa results before applying RF-RFE

In the second experiment, the RF-RFE feature selection algorithm was applied. There was a steady increase in accuracy for all the algorithms in the optimised dataset as shown in Table III and Fig 5. The accuracy of the GBMs algorithm increased by 2.42%. The accuracy of the FFNN algorithm also increased by 2.71 while the accuracy of the ELMs algorithm increased by only 0.1. The FFNN algorithm had the highest increment in the Kappa value. It increased by 5.44%, followed by the accuracy of the GBMs algorithm, which increased by 4.84%. The ELMs algorithm had the least increment in accuracy with a value of 0.22%.

TABLE III. RESULTS AFTER APPLYING RF-RFE

Method	Accuracy (%)	Kappa (%)
FFNN	99.66	99.33
ELMs	96.63	93.27
GBMs	100.00	100.00

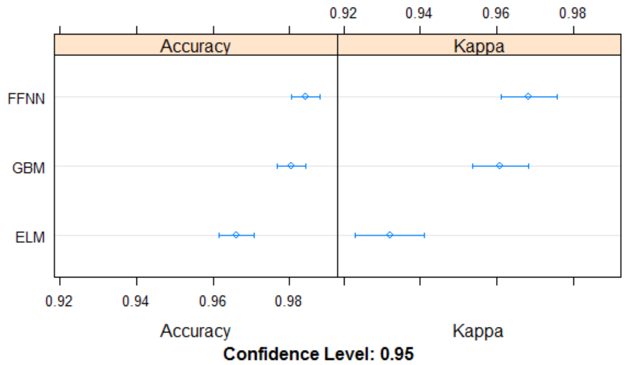

Fig. 5. Accuracy and Kappa results after applying RF-RFE

VI. CONCLUSION

In this paper, we have presented the RF-RFE feature selection for gender voice recognition. The results of the experiments cconducted in this study showed that the RF-RFE increases the gender voice recognition accuracy. However, the increment was small for all the three comparative algorithms that is the GBMs, FFNN and ELMs. There is a high probability that the increment can improve significantly on vast amounts of data. We have further discovered that the GBM algorithm is a technique that can be applied in gender voice recognition because it shockingly gave exceptionally high recognition accuracy. The FFNN and ELMs can also be used in this area because they achieved an accuracy, which were also relatively high. However, the GBMs and FFNN were slow in performing the computations when compared to the ELMs. If GBMs and FFNN algorithms were to be considered as candidates for the task of gender voice recognition, they would require bigger processing power. We would like to create a new dataset with our local languages to explore the performance of this technique and see how it performs on a different dataset as our future work.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interests regarding the publication of this paper.

REFERENCES

[1] M. Gupta, S. S. Bharti and S. Agarwal, "Support vector machine based gender identification using voiced speech frames," *2016 Fourth International Conference on Parallel, Distributed and Grid Computing (PDGC)*, Waknaghat, pp. 737-741, 2016.

[2] G.A Bryant, and H.C. Barrett,"Vocal emotion recognition across disparate cultures", *Journal of Cognition and Culture*, vol. 8, issue 1-2, pp. 135-148, 2008.

[3] R. B. Zajonc, "Felling and thinking: preferences need no inferences," *American Psychologist*, vol. 35, pp. 151-175, 1980.

[4] M. Fairhurst, M. Erbilek, and M. Da Costa-Abreu, "Selective review and analysis of aging effects in biometric system implementation", *IEEE Transactions on Human-Machine Systems*, vol. 45, issue 3, pp. 294-303, 2015.

[5] M. H. Bahari, M. McLaren, H. van Hamme, and D.A. van Leeuwen, "Speaker age estimation using i-vectors", *Engineering Applications of Artificial Intelligence*, vol. 34, pp. 99-108, 2014.

[6] H. Harb and L. Chen, "Gender classification using a general audio classifier", *IEEE International Conference on Multimedia & Expo*, vol. 1, pp. 733-736, 2003.

[7] M. Kotti and C. Kotropoulos, "Gender classification in two emotional speech database", *19th International Conference on Pattern Recognition* ,pp. 6-9,2008.

[8] H. Harb and L. Chen, "Voice-based gender identification in multimedia applications," *Journal of Intelligent Information Systems*, vol. 24, no. 2, pp. 179–198, 2005.

[9] Y. M. Zeng, Z. Y. Wu, T. Falk and W. Y. Chan, "Robust GMM Based Gender Classification using Pitch and RASTA-PLP Parameters of Speech," *2006 International Conference on Machine Learning and Cybernetics*, Dalian, China, pp. 3376-3379, 2006.

[10] W. Li, D. Kim, C. Kim, and K. Hong, "Voice-based recognition system for non-semantic information by language and gender", *3rd International Symposium on Electronic Commerce and Security*, pp. 84-88, 2010.

[11] M. Alhussein, Z. Ali, M. Imran, W. Abdul," Automatic gender detection based on characteristics of vocal folds form Mobile healthcare system", *Mobile Information Systems*, vol. 2016, 2016

[12] M. Hossain, G. Muhammad, M. Alhamid, B. Song, K. Al-Mutib, "Audio-visual emotion recognition using big data towards 5G", *Mobile Networks and Applications*, vol. 21, issue 5, pp. 753-763, 2016.

[13] D. Verma, D. Mukhopadhyay, and E. Mark, "Role of gender influence in vocal Hindi conversations: a study on speech emotion recognition", *International Conference on Computing Communication Control and automation (ICCUBEA)* , pp. 1-6, 2016.

[14] A. Aly, and A. Tapus," Towards an online voice-based gender and internal state detection model.", *Human-Robot Interaction (HRI), 6th ACM/IEEE International Conference*, pp. 105-106, 2011.

[15] J. Rong, Y. P. P. Chen, M. Chowdhury and G. Li, "Acoustic Features Extraction for Emotion Recognition," *6th IEEE/ACIS International Conference on Computer and Information Science (ICIS 2007)*, Melbourne, Qld., pp. 419-424, 2007.

[16] P. Castellano, S. Slomka, and P. Barger,"Gender gates for telephone-based automatic speaker recognition.", *Digital Signal Processing: A Review Journal*, vol. 7, issue 2, pp. 65-79, 1997.

[17] K. Wu and D. G. Childers," Gender recognition from speech. Part I: Coarse analysis", *The Journal of the Acoustical Society of America*, vol. 90, issue 4 pp. 1841-1856, 1991.

[18] T. Vogt and E. Andre,"Improving automatic emotion recognition from speech via gender differentiation", *Proc. Language Resources and Evaluation Conference*, pp. 1123-1126, 2006.

[19] H. Harb and L. Chen," A general audio classifier based on human perception motivated model", *Multimedia Tools and Applications*, vol. 34, issue 3, pp. 375-395, 2007.

[20] J.G. Wang, J. Li, W.-Y. Yau and E. Sung, "Boosting dense SIFT descriptors and shape contexts of face images for gender recognition", *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops*, pp. 96-102, 2010.

[21] M. Kumari, and N. Talukdar," A new gender detection algorithm considering the non-stationarity of speech signal", *2nd International Conference on Communication, Control and Intelligent Systems*, pp. 141-146, 2017.

[22] E. Jung, A. Schwarzbacher, and R. Lawlor, "Implementation of real-time AMDF pitch-detection for voice gender normalisation", *International Conference on Digital Signal Processing*, vol. 2 pp. 827-830, 2002.

[23] A. Sabir, N. Al-Jawad, S. Jassim, and A. Al-Talabani, "Human gait gender classification based on fusing spatio-temporal and wavelet statistical features.", *5th Computer Science and Electronic Engineering Conference*, pp. 140-145, 2013.

[24] M. Ichino, Y. Yamazak, W. Jian-Gang and Y. W. Yun," Text independent speaker gender recognition using lip movement.", 12th *International Conference on Control, Automation, Robotics and Vision*, pp. 176-181, 2012.

[25] T. Maka and P. Dziurzanski, "An analysis of the influence of acoustical adverse conditions on speaker gender identification," *XXII Annual Pacific Voice Conference (PVC)*, Krakow, pp. 1-4 , 2014.

[26] M. Kumari and I. Ali, "An efficient algorithm for gender detection using voice samples," *Communication,Control and Intelligent Systems (CCIS)*, pp. 221-226, 2015.

[27] G.S Archana, and M. Malleswari," Gender identification and performance analysis of speech signals.", *Proceedings of 2015 Global Conference on Communication Technologies* , pp. 483-489, 2015

[28] S. Gaikwad, B. Gawali, and S. C Mehrotra, "Gender identification using SVM with combination of MFCC.", *International Journal of Computer Science and Electronics Engineering (IJCSEE)*, vol. 3, issue 5, pp. 351-355, 2016.

[29] T.Bocklet, A. Maier, J. G. Bauer, F. Burkhardt, and E. Noth," Age and gender recognition for telephone applications based on GMM supervectors and support vector machines.", *ICASSP, IEEE International Conference on Acoustics*, pp. 1605-1608, 2008.

[30] M. Ichino, Y. Yamazak, W. Jian-Gang and Y. W. Yun,"Improved estimation of femininity using GMM supervectors and SVR for voice therapy of gender identity disorder clients", *ICASSP, IEEE International Conference on Acoustics*, pp. 7751-7754, 2013.

[31] E. Ramdinmawii and V. K. Mittal," Gender identification from speech signal by examining the speech production characteristics", *International Conference on Signal Processing and Communication*, pp. 244-249, 2017.

[32] R. Rajeswara Rao," Robust feature based automatic text-independent gender identification system using ergodic hidden Markov models (HMMs)" , *International Journal of Engineering and Computer Science*, Volume 3 Issue 8, pp. 7874-7878, 2014.

[33] E. Fokoue, and Z. Ma, "Speaker gender recognition via MFCCs and SVMs", 2013, [online] Available: http://scholarworks.rit.edu/article/1749.

[34] J. Přibil, A. Přibilová and J. Matoušek, "GMM-based speaker gender and age classification after voice conversion," *2016 First International Workshop on Sensing, Processing and Learning for Intelligent Machines (SPLINE)*, pp. 1-5, 2016.

[35] T. Jayasankar, K. Vinothkumar and Arputha Vijayaselvi, "Automatic gender identification in speech recognition by genetic algorithm.", *Applied Mathematics and Information Sciences*, vol. 11, issue 3, pp. 907-913, 2017.

[36] Dataset,https://raw.githubusercontent.com/primaryobjects/voicegender/master/voice.csv.

[37] M. Kuhn,"Variable selection using the caret package." Available: http://cran.r-project.org/web/packages/caret/vignettes/caretSelection.pdf. Accessed 7 May 2012.

[38] I. Guyon, J. Weston, S. Barnhill, and V. Vapnik, Gene selection for cancer classification using support vector machines, *Machine Learning*, vol.46, issue.1/3, pp.389-422, 2002.

[39] A. I. Galushkin, "Neural Networks Theory", Neural Networks, vol. 3, pp. 294- 302, 2001.

[40] M. H. Ali, "Design fast feed forward neural networks to solve two point boundary value problems," [M.S. Thesis], University of Baghdad, College of Education Ibn Al-Haitham, 2012.

[41] A. Natekin, and A. Knoll, "Gradient boosting machines atTutorial," *Frontiers in Neurorobotics*, vol. 7, pp. 1-21, 2013.

[42] J. Friedman, "Greedy function approximation: a gradient boosting machine", *Annals of Statistics*, vol. 29, issue 5, pp. 1189-1232, 2001.