International Journal of Engineering Applied Sciences and Technology, 2020
Vol. 4, Issue 11, ISSN No. 2455-2143, Pages 568-573
Published Online March 2020 in IJEAST (http://www.ijeast.com)

PERFORMANCE EVALUATION OF ALGORITHMS FOR GENDER CLASSIFICATION

Shivani Sanjay Tejale
BE Student
Department of Computer Engineering,
NDMVP'S college of Engineering, Nashik

Tushar B. Kute
Researcher
MITU Skillologies
Pune (MH), India-411027

Abstract— **Gender Classification based on the acoustic properties of voice sample is an important task for various applications. Gender of the person can be identified by the acoustic properties of voice. For classification choosing one of the best matching learning is another important task. In this proposed study we conducted various experiment study of various machine learning algorithms for Gender Classification. We analysed the performance of algorithms using voice dataset. From this we concluded that Random Forest and Artificial Neural Networks giving best results.**

Keywords—Machine learning; Kernel methods; Deep learning; SVM; Artificial Neural networks.

I. INTRODUCTION

Gender is the most significant characteristic of speech. Gender classification is a technique that aims to determine the gender of the speaker through speech signals analysis. Gender prediction is important in applications like targeted advertisements, interactive systems and mobile based health care systems. Based on the gender of a person interactive systems respond accordingly. If marketing firms know the gender of the person then they can target respective people who potentially buy the products. Classifying the gender of a person accurately based on their voice is a challenging problem in machine learning.

The "deep" in "Deep Learning" refers to the numbers of layers the data is transformed. It is a class of machine learning that use multiple layers to progressively extract higher level features from raw input. Deep learning models are more suitable for unstructured data like audio, video and images .For example in image processing ,lower level mat identify edges ,while higher level may identify human-meaningful items such as faces/letters or digits. Deep learning is being used today in our cell phones, cars, and tablets and computers. Deep learning models perform better results when the data is large.

In this paper we used the voice dataset consists of 3168 male and female voice acoustic features to train different machine learning algorithms. The dataset consists of the unique acoustic properties of the each voice of the respective gender which is required to train the various machine learning algorithms. From this research we compared the accuracy of different algorithms.

II. RELATED WORK

There are numerous machine learning, deep learning models to classify the person is male or female based on speech. In [1] with Support Vector Machines attained 95% accuracy for the gender classification system. In [2], pitch was used for the gender classification with Multi Layer Perception Neural networks chived the accuracy of 96%. In [3] Support Vector Machines, Classification and Regression Tree (CART) [4] models were used. In [5] Lee and Lang used Support Vector Machine(SVM). In [6] Silvosky and Nouza used Gaussian Mixture Models(GMM). In [7] by using Multilayer Perceptron (MLP) networks achived 96.74% accuracy.

III. SPEECH DATASET

The speech dataset [10] has 3168 voice samples of male and female. Each sample consists acoustic properties of voice.

Dataset file contains the following fields [9]:

meanfreq, mode, sd, centroid, Q25, Q75, skew, IQR, kurt, sp.ent, meanfun, minfun ,maxfu, mindom, meandom, maxdom, dfrange, modindex, label.

"label" field contains the values for male or female classification. 0 indicates the properties of male while 1 indicates the properties of female.

The remaining fields are acoustic properties of voice dataset described in TABLE I.

International Journal of Engineering Applied Sciences and Technology, 2020
Vol. 4, Issue 11, ISSN No. 2455-2143, Pages 568-573
Published Online March 2020 in IJEAST (http://www.ijeast.com)

TABLE I. ACOUSTIC PROPERTIES OF EACH VOICE SAMPLE

Acoustic Properties	
Properties	***Description***
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 (see note in specprop description)
kurt	kurtosis (see note in specprop description)
sp.ent	spectral entropy
sfm	spectral flatness
mode	mode frequency
centroid	frequency centroid
peakf	peak frequency (frequency with highest energy)
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. Calculated as the accumulated absolute difference between adjacent measurements of fundamental frequencies divided by the frequency range
label	male or female

IV. PERFORMANCE ANALYSYS OF ALGORITHMS

Various machine learning models exists for classifying the gender of the person. These include Logistic Regression, KNN, Naïve Bayes, Decision Tree, Random Forest, Support vector machine, Artificial neural networks Classification algorithms are used for solving problems like identification of person gender, intruder detection and Spam detection etc. In this research paper we compared classification algorithms using voice dataset.

We did conduct experiment with machine learning classification algorithms on voice dataset and observed the train and test set accuracies for seven classification algorithms. Amongst all these algorithms the results is measured according to the accuracies with the different algorithms setup.

We used sklearn pre-processing library for data pre-processing. The dataset is read as voice.csv file .In voice dataset no missing values present in the dataset. Initially we load the input data to the machine. Then the model is trained, connection to the input variable and output is made. Here in this study the output variable is a "label " field having binary numbers as 0 and 1,where 0 indicates male while 1 indicates female value. Machine learns the relationship between given input and output. The "label" field is then applied standard scalar for standardization of values . We used pandas, NumPy packages to load the dataset, to perform numerical calculations respectively and sklearn package used for modelling the machine learning algorithm. In all the experiments test set size is 0.25. Keras and TensorFlow used in Artificial Neural Networks(ANN). We used 10 fold cross validation to train the models. The accuracies are shown in the Table II.

Both the Decision Tree and Random Forest Algorithm are giving better results compared with other machine learning algorithms. Decision Tree is giving 95.5% with 0.04 secs while Random Forest is giving the same accuracy i.e. 95.5% but the prediction time is more compared to the Decision Tree Algorithm i.e. 0.14 secs. Hence we have used the Decision Tree Algorithm for Gender Classification. SVM is giving 97% accuracy on both train, test sets with linear kernel. Artificial Neural Network with three hidden dense layer of each contains 1000 nodes and relu as activation function, one input layer with 20 features and one output layer consists two nodes. In the

International Journal of Engineering Applied Sciences and Technology, 2020
Vol. 4, Issue 11, ISSN No. 2455-2143, Pages 568-573
Published Online March 2020 in IJEAST (http://www.ijeast.com)

output layer SoftMax is used as activation function and adam optimizer used then ANN is giving 98% accuracy.

From the TABLE II we can conclude Decision Tree and Random Forest have better accuracies. Parameters influences the machine learning algorithm performance. We can further improve these algorithms by parameter tuning.

TABLE II. ACCURACY OF MACHINE LEARNING MODELS

Accuracy(%)		
Model	***Training set***	***Time***
Logistic Regression	91.54	0.31
KNN	72.47	0.15
Naive Bayes	88.5	0.005
Decision Tree	95.5	0.040
Random Forest	95.5	0.143
SVM	74.24	0.38
ANN	93.9	1.54

V. PROPOSED WORK

I] Data Collection

The dataset for study can be downloaded from the Kaggle website.www.kaggle.com. Unzip the downloaded zip file and place the "voice.csv" file in your local drive. This is the file that we are going to use to train our machine learning model.

II] Data Pre-processing

The data is pre-processed by processing all the variables except the "label " field as the input instances to the machine learning algorithms whereas the "label " field is used as the output instance for the gender prediction. The 25% of the data is processed to test the machine and the remaining is used to train the machine .The input instances includes meanfreq, mode, sd, centroid, Q25, Q75, skew, IQR, kurt, sp.ent, meanfun, minfun,maxfu, mindom, meandom, maxdom, dfrange, modindex while the output include the label field.

The following figures shows the dataset fields.

dataset

	meanfreq	sd	median	Q25	Q75	IQR	skew	kurt	sp.ent	sfm	...	centroid	meanfun	minfun	maxfun
0	0.059781	0.064241	0.032027	0.015071	0.090193	0.075122	12.863462	274.402906	0.893369	0.491918	...	0.059781	0.084279	0.015702	0.275862
1	0.066009	0.067310	0.040229	0.019414	0.092666	0.073252	22.423285	634.613855	0.892193	0.513724	...	0.066009	0.107937	0.015826	0.250000
2	0.077316	0.083829	0.036718	0.008701	0.131908	0.123207	30.757155	1024.927705	0.846389	0.478905	...	0.077316	0.098706	0.015656	0.271186
3	0.151228	0.072111	0.158011	0.096582	0.207955	0.111374	1.232831	4.177296	0.963322	0.727232	...	0.151228	0.088965	0.017798	0.250000
4	0.135120	0.079146	0.124656	0.078720	0.206045	0.127325	1.101174	4.333713	0.971955	0.783568	...	0.135120	0.106398	0.016931	0.266667
5	0.132786	0.079557	0.119090	0.067958	0.209592	0.141634	1.932562	8.308895	0.963181	0.738307	...	0.132786	0.110132	0.017112	0.253968
6	0.150762	0.074463	0.160106	0.092899	0.205718	0.112819	1.530643	5.987498	0.967573	0.762638	...	0.150762	0.105945	0.026230	0.266667
7	0.160514	0.076767	0.144337	0.110532	0.231962	0.121430	1.397156	4.766611	0.959255	0.719858	...	0.160514	0.093052	0.017758	0.144144
8	0.142239	0.078018	0.138587	0.088206	0.208587	0.120381	1.099746	4.070284	0.970723	0.770992	...	0.142239	0.096729	0.017957	0.250000
9	0.134329	0.080350	0.121451	0.075580	0.201957	0.126377	1.190368	4.787310	0.975246	0.804505	...	0.134329	0.105881	0.019300	0.262295
10	0.157021	0.071943	0.168160	0.101430	0.216740	0.115310	0.979442	3.974223	0.965249	0.733693	...	0.157021	0.088894	0.022069	0.117647
11	0.138551	0.077054	0.127527	0.087314	0.202739	0.115426	1.626770	6.291365	0.966004	0.752042	...	0.138551	0.104199	0.019139	0.262295
12	0.137343	0.080877	0.124263	0.083145	0.209227	0.126082	1.378728	5.008952	0.963514	0.736150	...	0.137343	0.092644	0.016789	0.213333
13	0.181225	0.060042	0.190953	0.128839	0.229532	0.100693	1.369430	5.475600	0.937446	0.537080	...	0.181225	0.131504	0.025000	0.275862
14	0.183115	0.066982	0.191233	0.129149	0.240152	0.111004	3.568104	35.384748	0.940333	0.571394	...	0.183115	0.102799	0.020833	0.275862
15	0.174272	0.069411	0.190874	0.115602	0.228279	0.112677	4.485038	61.764908	0.950972	0.635199	...	0.174272	0.102046	0.018328	0.246154

dataset

IQR	skew	kurt	sp.ent	sfm	...	centroid	meanfun	minfun	maxfun	meandom	mindom	maxdom	dfrange	modindx	label
0.075122	12.863462	274.402906	0.893369	0.491918	...	0.059781	0.084279	0.015702	0.275862	0.007812	0.007812	0.007812	0.000000	0.000000	male
0.073252	22.423285	634.613855	0.892193	0.513724	...	0.066009	0.107937	0.015826	0.250000	0.009014	0.007812	0.054688	0.046875	0.052632	male
0.123207	30.757155	1024.927705	0.846389	0.478905	...	0.077316	0.098706	0.015656	0.271186	0.007990	0.007812	0.015625	0.007812	0.046512	male
0.111374	1.232831	4.177296	0.963322	0.727232	...	0.151228	0.088965	0.017798	0.250000	0.201497	0.007812	0.562500	0.554688	0.247119	male
0.127325	1.101174	4.333713	0.971955	0.783568	...	0.135120	0.106398	0.016931	0.266667	0.712812	0.007812	5.484375	5.476562	0.208274	male
0.141634	1.932562	8.308895	0.963181	0.738307	...	0.132786	0.110132	0.017112	0.253968	0.298222	0.007812	2.726562	2.718750	0.125160	male
0.112819	1.530643	5.987498	0.967573	0.762638	...	0.150762	0.105945	0.026230	0.266667	0.479620	0.007812	5.312500	5.304688	0.123992	male
0.121430	1.397156	4.766611	0.959255	0.719858	...	0.160514	0.093052	0.017758	0.144144	0.301339	0.007812	0.539062	0.531250	0.283937	male
0.120381	1.099746	4.070284	0.970723	0.770992	...	0.142239	0.096729	0.017957	0.250000	0.336476	0.007812	2.164062	2.156250	0.148272	male
0.126377	1.190368	4.787310	0.975246	0.804505	...	0.134329	0.105881	0.019300	0.262295	0.340365	0.015625	4.695312	4.679688	0.089920	male
0.115310	0.979442	3.974223	0.965249	0.733693	...	0.157021	0.088894	0.022069	0.117647	0.460227	0.007812	2.812500	2.804688	0.200000	male
0.115426	1.626770	6.291365	0.966004	0.752042	...	0.138551	0.104199	0.019139	0.262295	0.246094	0.007812	2.718750	2.710938	0.132351	male
0.126082	1.378728	5.008952	0.963514	0.736150	...	0.137343	0.092644	0.016789	0.213333	0.481671	0.015625	5.015625	5.000000	0.088500	male
0.100693	1.369430	5.475600	0.937446	0.537080	...	0.181225	0.131504	0.025000	0.275862	1.277114	0.007812	2.804688	2.796875	0.416550	male
0.111004	3.568104	35.384748	0.940333	0.571394	...	0.183115	0.102799	0.020833	0.275862	1.245739	0.203125	6.742188	6.539062	0.139332	male
0.112677	4.485038	61.764908	0.950972	0.635199	...	0.174272	0.102046	0.018328	0.246154	1.621299	0.007812	7.000000	6.992188	0.209311	male

Fig1. Dataset Fields

III] Train and Test Data Sets

Once the dataset is processed, we need to divide it into two parts. Split the dataset into Train and Test set. We will import and use the train_test_split()function for that. All variables except "label" will be the input values x for the Machine learning models. The variable "label" will be stored in y_predict, and will represent the prediction variable. We arbitrarily chose to allocate 25% of the total data for the training set.

IV] Overview of Different Algorithms.

a) Logistic Regression

Logistic Regression measures the relationship between the dependent variable which is categorical with one or more than one independent variable by estimating probabilities using a logistic function.it can generally be used where the dependent variable is binary that means the dependent variable can take only possible values like "yes" or "no", "0 " or "1".

b) KNN

KNN algorithm is one of most used algorithm in classification as it is a straight forward algorithm. A majority vote of an object is classified by its neighbours, with the purpose being assigned to the class most common among its l nearest neighbours. It can be used for regression -output is

International Journal of Engineering Applied Sciences and Technology, 2020
Vol. 4, Issue 11, ISSN No. 2455-2143, Pages 568-573
Published Online March 2020 in IJEAST (http://www.ijeast.com)

the value of the object. This value is the average of the benefit of its k nearest neighbours.

c)Naïve Bayes

Naïve Bayes is a classification algorithm which is based on the Bayes Theorem with an assumption of independence among predictors. In simple terms Naïve Bayes classifier assumes that a particular feature in a class is unrelated to any of the function in the class. This model is particularly used to build and extensively use the datasets.

d)Decision Tree

Decision Tree makes classification models in the form of tree structure. It breaks the large dataset into smaller subsets. An associates decision tree developed at the same time. The final result is a decision tree with the decision nodes and leaf nodes. A decision node as two or more branches. Leaf node represents the decision. The first decision node Tree corresponds to the best predictor called root node.

e) Random Forest

Random Forest is a supervised Learning algorithm. It creates a forest and makes it some how casual. It is one of the most accurate learning algorithms available. For larger datasets it produces a highly accurate classifier. Random forest is an effective method for estimating missing data and maintain accuracy when a larger amount of data is missing.

f) Support Vector Machine

Support vector machine is a linear model for classification problems creates a line or a hyperplane which separates the data into classes. So initially the first task of SVM is to find an ideal line that separates this dataset in two classes. But if there isn't a unique line finding the best line accordingly is the main task which separates the dataset. The main goal of Support Vector Machine is to get a generalized separator.

g) Artificial Neural Network

Artificial neural networks (ANNs) are able to learn something about what they see and then can generalize that knowledge to examples (or samples) that they have never seen before [1]. This is a very powerful capability that humans often take for granted because our brains automatically do it so well. You are able to understand the concept of a rock after seeing and perhaps touching very few examples of rocks. From that point on, you can identify any rock, even those that are shaped differently or have different colors or textures from the rocks you've seen before. This approach is very different from the traditional method of teaching or explicitly programming computers based on detailed rules that must cover every possible outcome.

Proposed GUI

The proposed GUI For Gender classification contains the input instances. These input variable are the labels .ie the input fields followed by the input from the user i.e. the spin box with the specified ranges of the each voice sample data. The range for the each acoustic properties of the voice is analysed from the dataset with the minimum and maximum value of each column so that when the user enters the input the entered values are checked with the specified range from the dataset values .The result is then predicted through these values that whether the entered values exists for a male or female. Prediction is done via a Button which predicts whether the person is a Male or Female.

The following figure shows the Overall GUI for Gender Classification.

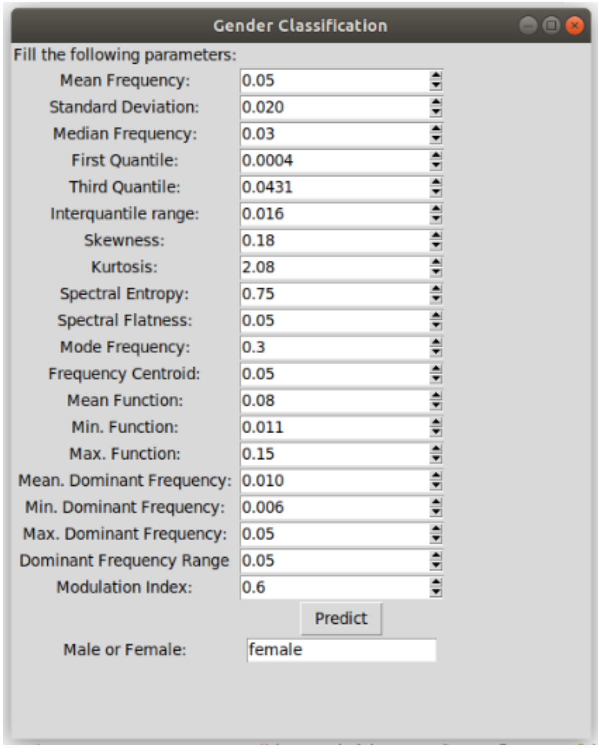

Fig2. Overall GUI of the System

Parameter tuning [11] is used to find the best hyper parameters. Grid Search technique is used to find best hyper parameters. Grid Search will test several combinations of hyper parameters and returns the best selection that gives best accuracy.

We created dictionary with hyper parameters and applied on GridSearchCV of karas library. GridSeachCV will train Artificial Neural Networks using k-fold cross validation to get relevant accuracy with different combinations of the dictionary of hyper parameters and returns best accuracy with best selection of these values.

International Journal of Engineering Applied Sciences and Technology, 2020
Vol. 4, Issue 11, ISSN No. 2455-2143, Pages 568-573
Published Online March 2020 in IJEAST (http://www.ijeast.com)

We applied parameter tuning on SVM using GridSeachCV with different kernels linear, rbf, poly, different gamma and C vales. Best parameter values are C 0.6, gamma 0.04 and kernel rbf. After applying parameter tuning SVM, ANN are giving improved results. We applied 0.1 dropout between hidden layer to avoid over fitting machine learning model.

VI. RESULT ANALYSIS

According to the survey the overall analysis of various algorithm is studied. As 25% of the data is used for testing and remaining 75% is used to train the model. Here are the accuracies of these algorithms.

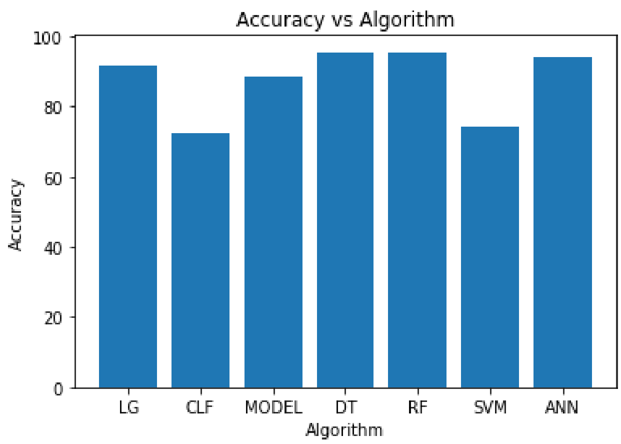

Fig3. Accuracies of various Algorithm

The Final result is calculated with the lowest time in which the algorithms predicts and the highest accuracy it gives .Following figure shows the prediction time of these algorithms.

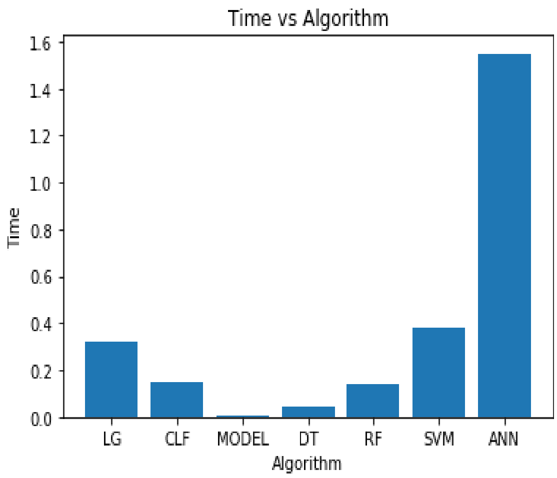

Fig4. Prediction Time of various Algorithm.

Hence analysing the above graphs it is concluded that the Random Forest and Decision Tree has same accuracy but the prediction time of Random Forest is more compared to Decision Tree Algorithms. So the Gender Classification id predicted by Decision Tree Algorithm.

The improved accuracies of Decision Tree, Random Forest are shown in the Table III.

TABLE III. ACCURACY OF MODELS AFTER PARAMETER TUNING

Accuracy(%)		
Model	***Training set***	***Time***
Decision Tree	95.5	0.040
Random Forest	95.5	0.143

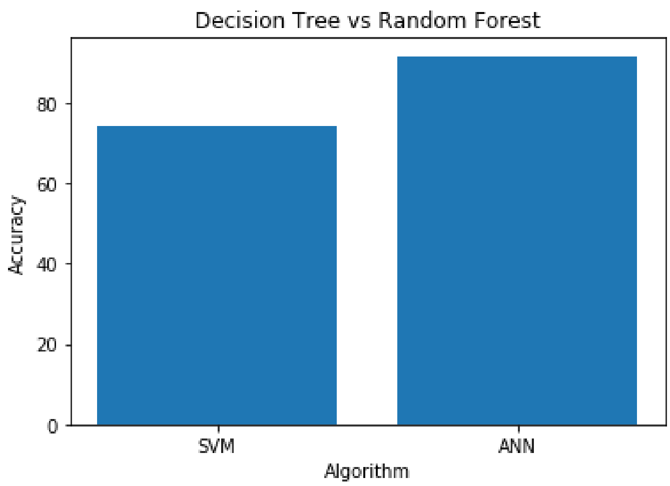

Fig5. Accuracy of Decision Tree vs Random Forest

Results of algorithms are shown in the Fig. 1. X-axis shows algorithms and Y-axis shows Accuracy of the algorithm.

VII. CONCLUSION

Random Forest and Decision Tree are performing better on voice dataset. Decision Tree is giving the 95.5% accuracy with 0.040 secs and Random Forest is giving 95.5% with 0.143 secs. From the above results we can conclude that Decision Tree Classifier models are performing better compared with machine learning algorithms to classify gender of a person using acoustic properties of voice.

VIII. REFERENCES

[1] S. Gaikwad, B. Gawali, and S. C. Mehrotra.(2012). Gender identification using SVM with combination of MFCC - Advances in Computational Research .vol. 4. no. 1,(pp. 69–73) .

[2] S. M. R. Azghadi, M. R. Bonyadi, and H. Sliahhosseini(2007). Gender classification based on feedforward backpropagation neural network - in Artificial Intelligence and Innovations : From Theory to Applications: Proceedings of the 4th IFIP International Conference on Artificial Intelligence Applications and Innovations (AIAI 2007), C.

[3] Cortes, Corinna & Vapnik, VN. (1995). Support Vector Networks. Machine Learning. 20. 273-297. 10.1023/A:1022627411411.

[4] L. Breiman, J.H. Friedman, R.A. Olshen, C.J(2008). Stone, Classification and Regression Trees, CRC Press.

[5] K.-H. Lee, S.-I. Kang, D.-H. Kim et al.(2008). A Support Vector Machine- Based Gender Identification Using Speech Signal. IEICE Transactions on Communications ,(pp. 3326-3329).

[6] J. Silovsky, and J. Nouza. Speech, Speaker and Speaker's Gender

Identification in Automatically Processed Broadcast Stream.

[7] Büyükyılmaz, Mücahit & Çıbıkdiken, Ali. (2016). Voice Gender Recognition Using Deep Learning. 10.2991/msota-16.2016.90.

[8] https://github.com/primaryobjects/voice-gender

[9] M. Araya-Salas, G. Smith-Vidaurre, warbler (2016): an R package to streamline analysis of animal acoustic signals. Methods Ecol Evolution, 2016, doi: 10.1111/2041-210X.12624.

[10] dataset, https://github.com/primaryobjects/voice- gender/blob/master/ voice.csv.

[11] G. S. Archana and M. Malleswari(2015). Gender identification and performanceAnalysis of speech signals - *2015 Global Conference on Communication Technologies (GCCT)*, Thuckalay, (pp. 483-489).

[12] M. A. Keyvanrad and M. M. Homayounpour.(2010) .Improvement on Automatic speaker gender identification using classifier fusion - *2010 18th IranianConference on Electrical Engineering*, Isfahan, Iran, (pp. 538-541).