Applied Acoustics 156 (2019) 351–358

Contents lists available at ScienceDirect

Applied Acoustics

journal homepage: www.elsevier.com/locate/apacoust

An effective gender recognition approach using voice data via deeper LSTM networks

Fatih Ertam

Department of Digital Forensics Engineering, Technology Faculty, Firat University, Elazig, Turkey

ARTICLE INFO

Article history:
Received 17 April 2019
Received in revised form 30 June 2019
Accepted 24 July 2019
Available online 1 August 2019

Keywords:
Gender recognition
Gender classification
Deep learning
Deeper LSTM
Machine learning

ABSTRACT

It is not difficult to estimate the gender of the human from other people's audio files. In general, people can easily identify the gender of the owner of a conversation with the experience they have acquired. However, it is not easy to predict whether a person is a man or a woman by computer systems. Hence, many papers and proposals have been presented to solve this problem using computer systems. In this study, Deeper Long Short Term Memory (LSTM) Networks structure was used for the prediction of gender from an audio data set. The study was successful at predicting gender with an accuracy of 98.4%. The proposed approach consists of 3 main steps. Firstly, 10 most effective data attributes were selected (i). Then, a deep learning-based network was created with the double-layer LSTM structure (ii). In addition to the performance comparison of the classification, accuracy values, sensitivity, and specificity performance metrics were also calculated (iii). At the same time, the accuracy of the proposed method was compared with the accuracy values obtained from the classifiers generated by conventional machine learning approaches. The study was successful at predicting gender with 98.4% success rate. It is thought that the study will be a pioneer in this field as an effective and fast approach for gender recognition.

1. Introduction and related works:

In order to implement a gender-based model, gender must be defined correctly. Defining the gender of the speaker is an old topic. However, the realization of this process using computer systems has begun to progress and develop in recent years [1]. Gender recognition is accepted as the removal of gender information from the speaker's speech. In general, it focuses on identifying a person from sound characteristics [2]. The classification of the speaker's gender information remains one of the most challenging problems in speech processing. Although many studies have been conducted focusing on feature extraction and classifier for improvement, classification accuracy is still not at the desired level. The key issue in determining the gender of the speaker is to produce robust features and to design a good classifier [3]. Computer systems are widely used in areas such as health systems, education systems and law, and systems are automated. A healthcare computer system can offer more useful options to a person whose gender can be determined [4,5]. Various studies focusing on gender recognition from sound files are found in the literature [6]. Li et al. [1] In their study, it was shown that the combination of various methods together with the Support Vector Machine would cause a low calculation cost. A combination of different number of systems is also used at the score level. Each system has complementary information from other systems. Approximately 53% accuracy has been achieved by combining different systems. Metze et al. [7] studied different techniques for gender classification based on telephone applications. They compared the performance of the systems they prepared with human listeners. They tried to deal with automatic voice recognition and language identification problems in their first systems called PPR. The core of this system is to create a PPR for each class in the gender database. They reported that the PPR system works almost like a human listener, and has the disadvantage of losing quality and accuracy in short phrases. Their second technique is based on prosodic properties. Various properties are used in this technique. in this technique uses statistical information on vibration, sparkle, noise ratio of harmonics, and many statistical information of basic frequency. All these properties were made and analyzed using a two-layer system. The first layer analyzes properties using three different neural networks. The second layer dynamically processes the output information produced by the first layer using the Bayesian network. The system based on the prosodic properties showed better performance in changing the expression time. Their third technique is linear estimation analysis. Gaussian's distance distribution is considered to contain useful information about the age and gender of the

E-mail address: fatih.ertam@firat.edu.tr

https://doi.org/10.1016/j.apacoust.2019.07.033
0003-682X/

Table 1
Attributes of the used dataset.

ID	Features	Weight Ranking	Weight	Description
1	meanfreq.	20	– 0.00029217142912	mean freq.
2	sd	3	0.0251049737965184	standard deviation of freq.
3	median	6	0.0183009572132704	median freq.
4	Q25	5	0.0204033845899695	Quantile (1.)
5	Q75	8	0.0165723868833801	Quantile (3.)
6	IQR	2	0.0328882207940331	interquantile range
7	Skew	11	0.0077717379887284	skewness
8	kurt	16	0.0033542203541026	kurtosis
9	sp.ent	17	0.0032758438388387	spectral entropy
10	sfm	4	0.0224875564897115	spectral flatness
11	mode	7	0.0180443914753449	mode freq.
12	centroid	10	0.0096595326154853	frequency centroid
13	meanfun	1	0.0920360982503833	fundamental freq. (average)
14	minfun	18	0.0011230913682374	fundamental freq. (minimum)
15	maxfun	19	0.0008470130510325	fundamental freq. (maximum)
16	meandom	9	0.0111018537226274	Dominant freq. (average)
17	mindom	13	0.0061865677824917	Dominant freq. (minimum)
18	maxdom	14	0.0048296464005949	Dominant freq. (maximum)
19	dfrange	15	0.0047568764246090	range of dominant freq.
20	modindx	12	0.0067696193607656	modulation index
21	label			Class, Female (0) or male (1)

speaker. Although this system was successful in predicting gender, it was not successful in predicting age. Because young and adult speakers have almost the same Gaussian distribution. Konig and Morgan used Multilayer Perceptron (MLP) for gender determination [8]. DARPA has reached up 84% accuracy on resource management database. This dataset is a data obtained from 160 speakers speaking American English. One of the first studies conducted in this field is the study of Acero and Huang researchers [9]. They have achieved a 30% reduction in error rate using Hidden Markov Model. They often proposed a maximum likelihood approach independently of speech. Neti and Roukos used a simple pattern-matching approach using the Euclidean distance between male

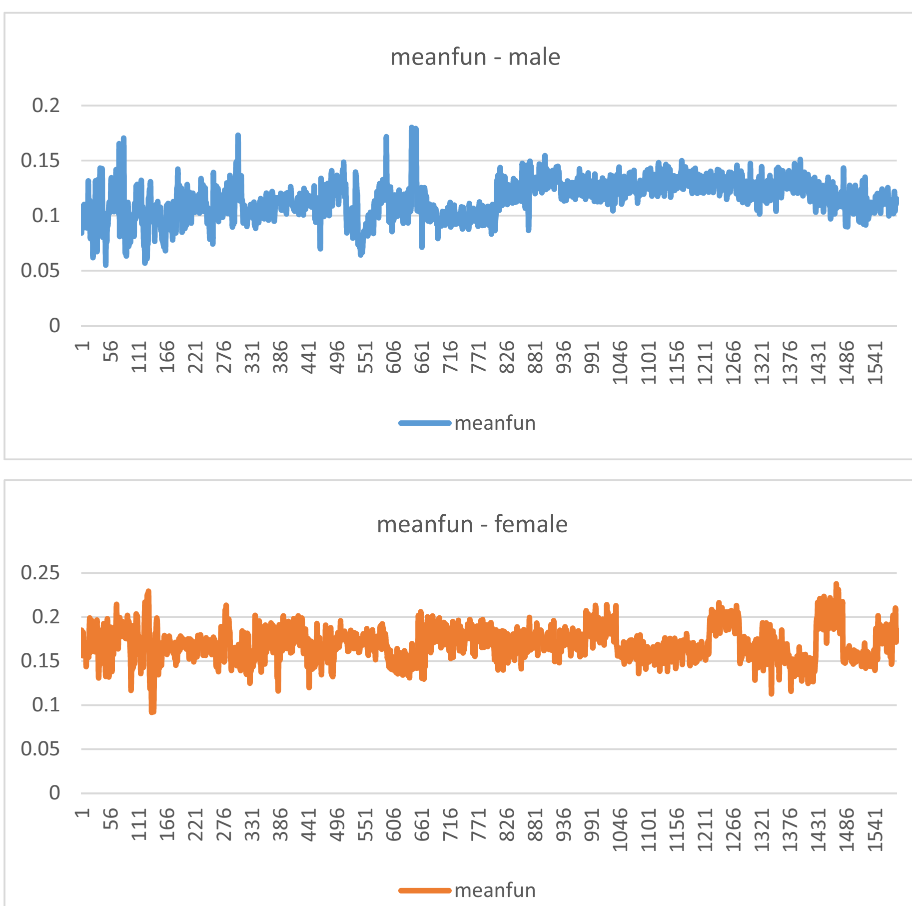

Fig. 1. Meanfun features acoustic parameters.

and female models. In this way, the shortest distance model was chosen to determine the gender of the person who spoke [10]. Ramadhan et al. used the random forest algorithm and it has been used in order to classified data for the data set used in the study by using parameter optimization. They was achieved 96.7% performance [11]. Harb et al. [12] proposed a new gender-based approach based on a general sound classifier. The method has been used to classify the language independently. They achieved 92% performance. There are various studies that attempt to estimate gender by using both face and voice data together [13]. Zorumand et al.[14] in their study they tried to identify the gender of children's voices. They examined the use of basic and formant frequencies to differentiate the sex of Malay children aged 7–12. In their study, they achieved a performance of 99.8% with MLP. Different studies with children's voices are also available in the literature. There are studies that provide pitch-range characteristics for age and gender classification [15]. Pahwa et al. [16] proposed a model using support vector machine and neurol network classifier. In their study, gender recognition system was formed with the samples taken from 46 speakers and reached to approximately accuracy of 93.5%. Buyukyilmaz et al. [17] used a multilayer perceptron deep learning model in order to define gender from their sound characteristics. The classifier model achieved accuracy performance of 96.8%. Zvarevashe et al. [18] using a data set consisting of 1584 male and 1584 female voices, using the gradient boosting machine algorithm and the random forest classifier were successful in identifying gender with a rate of 97.58% accuracy.

A deeper LSTM network was created for this study. The deep LSTM network was deepened by adding a new LSTM layer prior to the LSTM layer. LSTM is a gradient based model developed to solve the problem of backflow during the back-spreading of repetitive networks [19,20]. LSTM is a special type of Recurrent Neural Network (RNN) that can learn long-term dependencies [21–23].

The novel contributions of this work are as follow:

- Double-layer deep learning network structure was established and fast and high accuracy rate classification was performed.
- The performance of the established deep learning infrastructure compared to traditional classifiers was demonstrated.
- Instead of all features, the most weighted 10 features are selected. The first 10 features are provided to the created network.

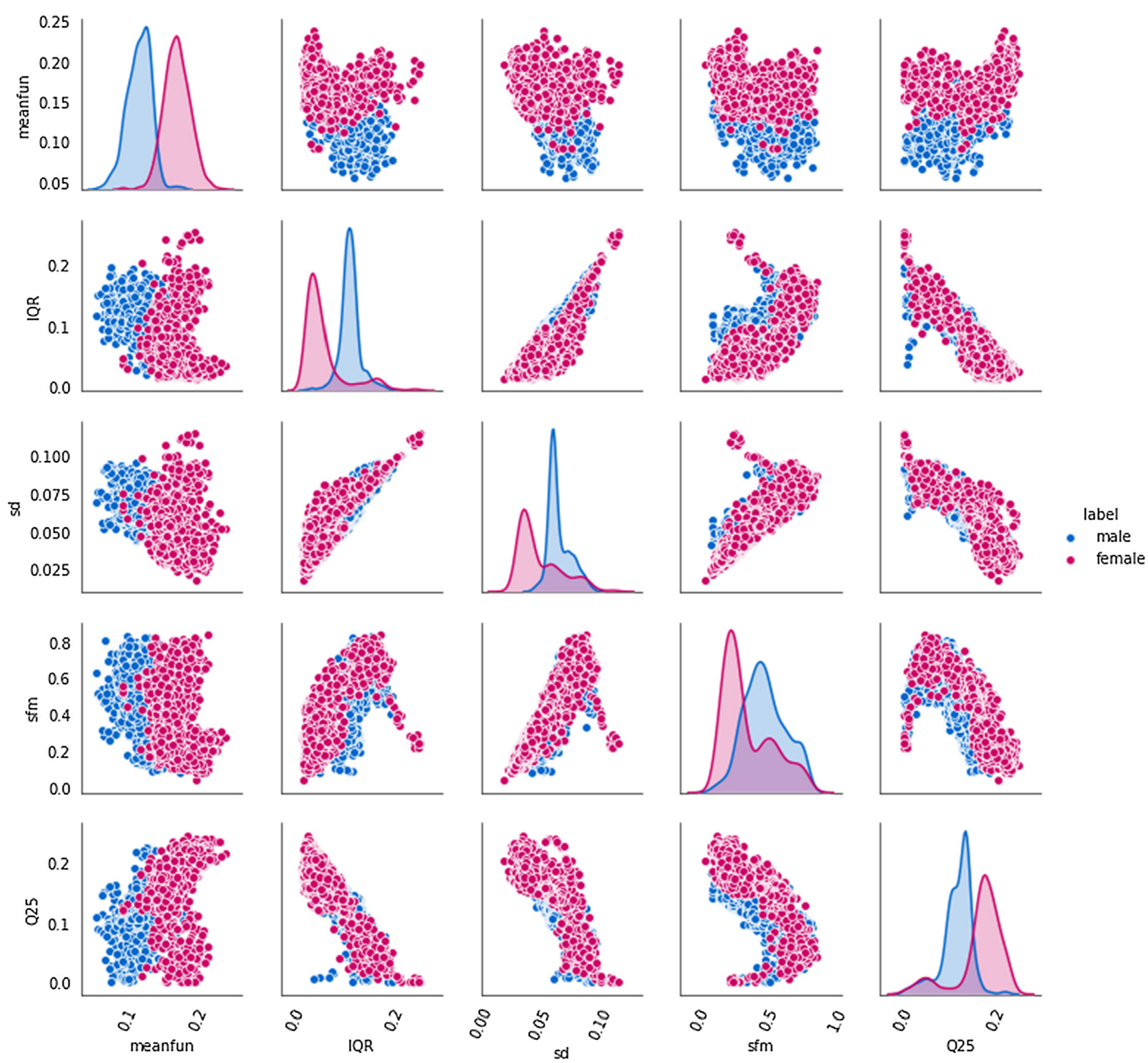

Fig. 2. Plot data (meanfun, IQR, sd, sfm, Q25).

The disadvantage of the study is that a small data set is used, which results in a lower performance increase compared to conventional machine learning approaches. The rest of the paper is structured as follows. The data set used in the second part of the study and the selected features are given. In the third chapter, the proposed method is given. In the last chapter, there are conclusions and discussions.

2. Materials

A public voice data set was used for this study [24]. This dataset was used to classify various machine learning approaches [17,25]. There are 3168 data in the data set. Half of the data set is marked as male and half as female voice. There are 20 features of this data. The attributes of this dataset are listed in Table 1. Table 1 also has the weight value and ranking of each features.

The best weight effect has been determined as feature *meanfun*. Fig. 1 shows the male and female acoustic parameters for the *meanfun* feature.

The Relieff-based feature selection method was used to calculate the weights. Relieff is an extension of the Relief algorithm, first proposed by Kira [26,27] and dealing with multi-category topics instead of only two types of data types [28]. When the Table 1 is examined, it is seen that the highest weight among the feautures belongs to the *meanfun* attribute. And the least effective effect belongs to the *meanfreq* features10 features with the highest weight for this study. These features are respectively *meanfun*, *IQR*, *sd*, *sfm*, *Q25*, *median*, *mode*, *Q75*, *meandom* and *centroid*. The plotting according to these 10 features is shown in Figs. 2 and 3.

On the histogram graph, the classes labeled as males are shown in blue, while the female-labeled classes are shown in red. When the figures are examined, it is seen that the histogram shows how male and female genders do not interfere with each other visually. The first 10 features selected in the histogram are given. In particular, it is observed that the gender of the female and male genders are clearly differentiated by the mean weight of the mean. In histogram curves, the values of these 10 features are shown visually by comparing with each other. The separation states of the sexes visually indicate the attributes to be chosen to differentiate between sexes. Thus, the performance of the classifier to be used can be predicted to be better by choosing which features.

3. Proposed methods

The proposed system consists of three stages: the reduction of properties, construction the deeper LSTM and testing the constructed deeper LSTM networks. The flow diagram of the proposed method is shown in Fig. 4.

In this study, a voice data set which is commonly used and labeled as male or female is used. In this data set, contains 21 attributes, one of which is used as a class. In order to increase the performance of the classifier, weight values were calculated for the 20 attributes other than the class. Relieff method was used in order to calculate weight values and to select desired attributes. Relieff is the development of the previously developed relief algorithm and the filter type is used as the feature selection algorithm. By using the Relieff method, features with a weight less than a certain value can be removed. In this study, half of the best weights were used and the other half was removed. Hence it was kept in the 10 feature dataset with the best weight. The features used are respectively: meanfun, IQR, sd, sfm, Q25, median, mode, Q75, meandom and centroid. The features removed are respectively: skew, modindx, mindom, maxdom, dfrange, wolf, sp.ent, minfun, maxfun and meanfreq. In the second stage, it is determined how much of the data set will be used for training and testing. For this purpose, the data set is randomly divided into 4 equal parts. 3 parts were used for training and 1 part for testing. In the 3rd stage, a deep learning-based classifier designed for training and testing of data was used. In this study, LSTM was used for deep learning based classifier. LSTM was developed from RNN. LSTM is a repetitive neural network architecture used in the field of deep learning. There is no feedback on standard forward-fed neural networks, but LSTM can successfully classify the entire data set and make a successful classification by means of this feedback and memory.

In order to increase the performance of the deep learning structure used, a two-layer LSTM architecture with a deeper structure was used. To this end, the LSTM network can be deepened by adding extra LSTM layers to the output mode before the LSTM layer. For this, the previous LSTM layer output mode is set to "*sequence*" and the last LSTM layer output mode to "*last*". Thus, LSTM-LSTM classifier was created. For both LSTM layers, the number of hidden neurons was chosen to be 100. The classifier "*epoch*" value was selected at 300. Performance metrics have been created by creating a confusion matrix.

The softmax activation function has been used in the network architecture used for deep learning. The softmax function produces a probability-based loss value using the score values generated by the artificial neural network. It has been applied to the test data by means of the ability gained by the classifier and it has been observed how successful it can make the gender prediction by making deeper LSTM classification.

The algorithm for the study is given below.

Algorithm 1: Pseudo code of the voice data classification.
Input: Voice Dataset (Train 75%, Test 25%)
Output: Male or Female Classification
1: Load attributes of the dataset
2: $[rank, weigths] = relieff(features, target)$; // Apply relief to calculate features ranks and weights
2: **for** i = 1 to 10 **do**
3: $features^{new}(:, i) = features(:, rank(i))$; // Select top 10 features from feature set.
4: **end for i**
5: Use hold out validation. 75% of the data are utilized as training and 25% of them are utilized for testing.
6: $acc^{mean} = 0$; // Define accuracy value
6: **for** j = 1 to 300 **do**
7: Train and test the data using LSTM-LSTM deeper networks. Obtain prediction values.
8: $acc = 0$;
9: **for** i = 1 to L **do** // L expresses 25% of the data.
10: **if** $prediction(i) = response(i)$ **then**
11: $acc = acc + 1$;
12: **end if**
13: **end for i**
14: $acc^{mean} = acc^{mean} + \frac{acc}{L}$;
15: **end for j**

---

4. Results and discussions

For classification, 75% of the data set of 3168 data was used for training. The rest were used for testing. Average values were obtained by running each classifier 300 times. For the study, a server computer with 32 GB RAM and Intel Xeon 2.16 GHz processor was used. Various classifiers have been created in order to compare the proposed Deeper LSTM approach with traditional machine learning approaches. These classifiers are explained in below.

Classifiers 1: Fine Tree is make use of as classifier.

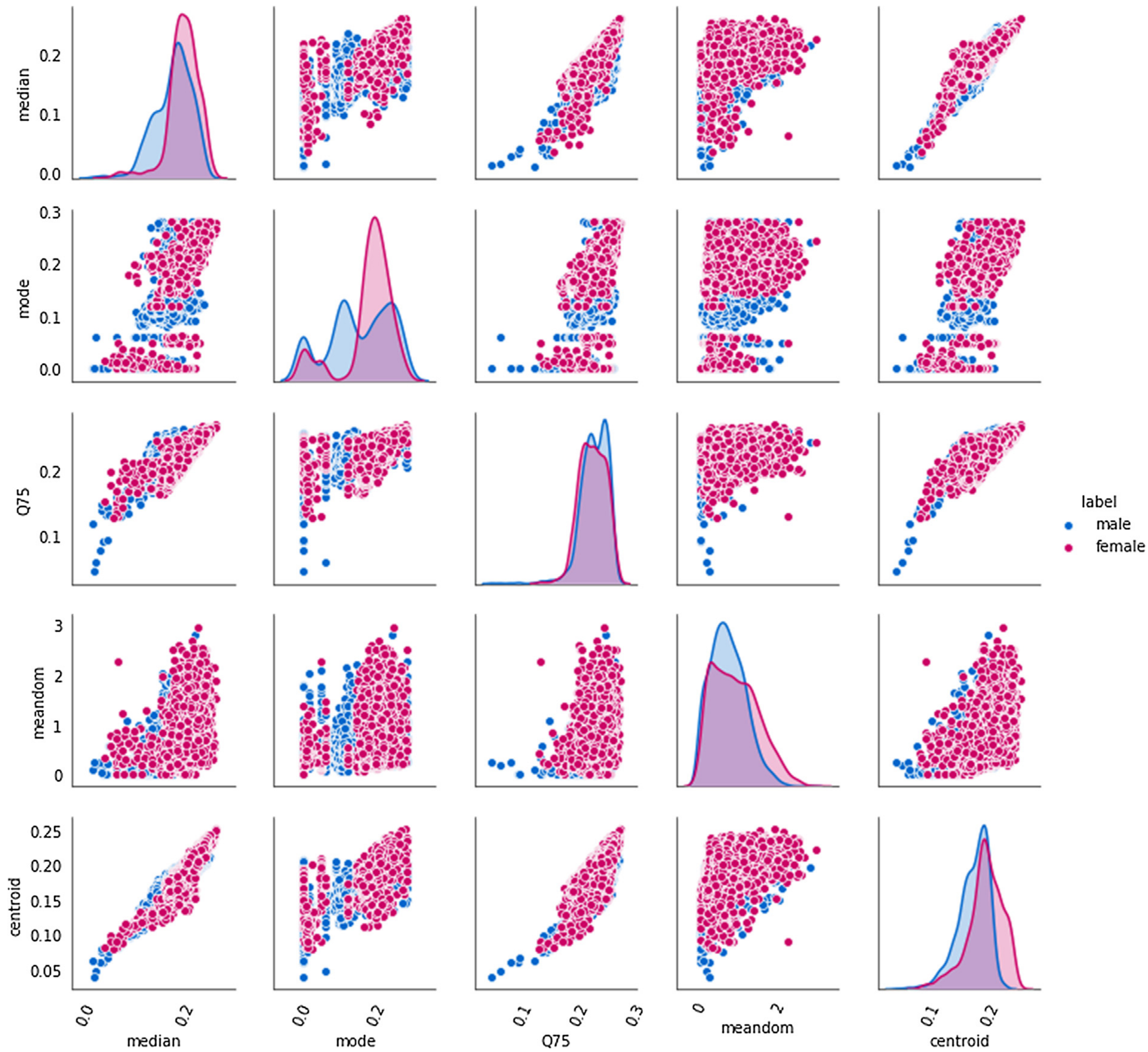

Fig. 3. Plot data (median, mode, Q75, meandom, centroid).

Classifiers 2: Linear Discriminant is make use of as classifier.

Classifiers 3: Logistic Regression is make use of as classifier.

Classifiers 4: Linear kernel Support Vector Machine (SVM) is make use of as classifier.

Classifiers 5: Quadratic kernel SVM is make use of as classifier.

Classifiers 6: Fine Gaussian kernel SVM is make use of as classifier.

Classifiers 7: Fine k nearest neighbor (kNN) is make use of as classifier.

Classifiers 8: Proposed Method; Deeper LSTM

Fine tree classifier; It is a widely used classifier especially in data mining and machine learning applications. The aim is to predict what the target might be based on some input variables [29].

The property of Linear Discriminant is that it models the dispersion of predictors severally in every of the response classes, and then it uses Bayes' theorem to prediction the probability [30].

Logistic Regression is a classification approach that is efficient for conditions where variables do not ever consist of quantitative values. Primary target is to define the possibility of acquisition another addicted variable by utilization detached variables [31].

kNN method is a popular classification method in data mining and machine learning because of its simplistic enforcement and remarkable classification performance [32].

SVM classification model has a quite widespread use within machine learning algorithms due to its superior level of generalization feature. This method has been often used and succeeded in the classification of supervised data [33]. The algorithm used in SVM is used in classification problems because it enables the statistical analysis of the data and the extraction of pattern on these data. The algorithm used forms a plane that bisects the qualifications appearing essentially at the size of the data set and a space with higher dimension best properly. Its use in classification problems is performed by transferring new samples to the space formed by the algorithm and trying to make classification by looking at their positions by the plane in this space [34].

performance metrics that are frequently used in the literature are used to evaluate the performance of the above mentioned classifiers. The confusion matrix is used to calculate these metrics. Table 2 also shows the confusion matrix used for the purpose of two-class data.

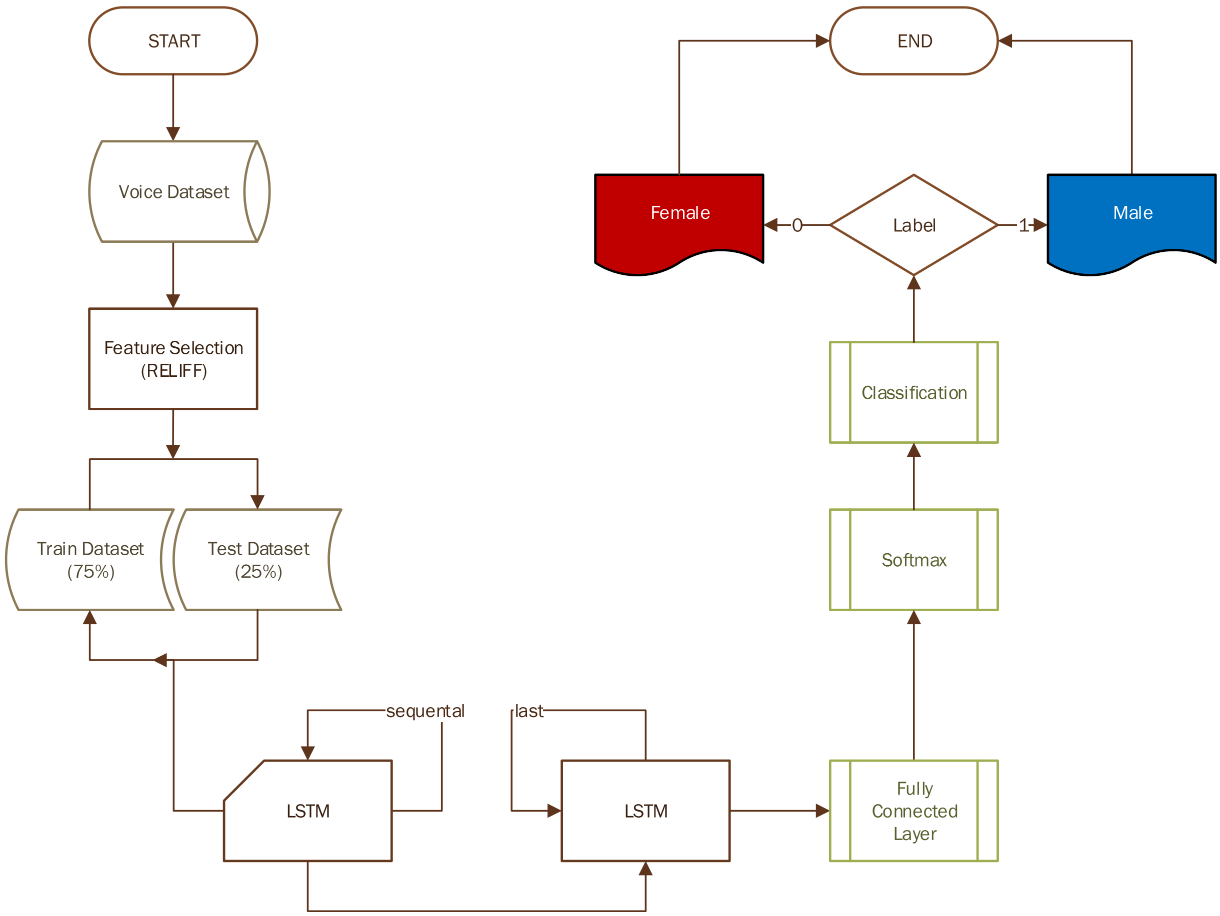

Fig. 4. Flow diagram of the proposed system.

In this study, as well as accuracy values, different performance metrics such as sensitivity and specificity were compared. Moreover, in this study, three other metrics are used, which are sensitivity (Sens), specificity (Spec), and the geometric mean (Gmean) of Sens and Spec. The metric Gmean is added to stabilize the accuracies of the 2 classes and so it is seen as the primary metric in our class imbalanced tasks. Accuracy (Acc) and The three employed metrics are de fined by Eqs. (1)–(4), in which the meaning of TP, FN, FP, and TN are expressed in Table 2.

$$Acc = \frac{TP + TN}{TP + FN + FP + TN} \tag{1}$$

$$Sens = \frac{TP}{TP + FN} \tag{2}$$

$$Spec = \frac{TN}{FP + TN} \tag{3}$$

$$Gmean = \sqrt{Sens \cdot Spec} \tag{4}$$

Table 2
Confusion matrix for gender recognition problem.

	Actual Values	
Predicted Values	True Positive (TP)	False Negative (FN)
	False Positive (FP)	True Negative (TN)

Table 3 demonstrates the results of the defined classifiers using accuracy, sensitivity, specificity and geometric mean. The comparisons of the classifiers are also listed in Table 3.

When the table is examined, it is seen that the deep learning approach gives better results in terms of both mean accuracy and geometric mean value according to classical machine learning approaches. It is observed that the fine tree classifier of classical machine learning approaches is worse than other classifiers. With the proposed method, it was effective to use a deeper design with double layer and to keep the information in memory with LSTM. The reason why all classifiers have achieved over 96% is related to the selection of attributes with the best weights by Relieff method. In the study, not only average accuracy values but also sensitivity and specificity values were calculated. Thus, the actual performance of the classifiers has been demonstrated by using different performance metrics. The linear discriminant and kNN reached the same results as the proposed method when the Sensitivity performance metric was examined. When the specificity performance value is examined, the proposed method has reached a very good result. classical machine learning approaches have lagged behind the proposed method. Instead of just taking one of the sentivity or specificity values, the geometric mean values have been subtracted from the sensitivity and specificity performance metrics. The proposed method for the geometric mean value was more successful than the other classifiers.

There are several studies using the same voice dataset in the literature. The comparison of the proposed method with the literature studies is given in Table 4.

Table 3
Results of the defined classifiers (%).

Classifier	Acc	Sens	Spec	Gmean
Fine Tree	96.2	94.6	97.9	96.2
Linear Discriminant	96.6	**97.4**	95.7	96.5
Logistic Regression	96.5	96.0	96.9	96.4
SVM (Linear)	96.3	96.7	96.6	96.6
SVM (Quadratic)	97.2	97.2	97.2	97.2
SVM (Gaussian)	97.3	96.9	97.7	97.3
kNN	97.6	**97.4**	97.7	97.5
Proposed Method	**98.4**	97.2	**99.5**	**98.3**

Table 4
Comparison of proposed method with the literature works.

Studies	Algorithm	Acc (%)
Ref [6]	Sequential minimum optimization	98.0
Ref [6]	kNearest neighbor	97.5
Ref [6]	C4.5 decision tree algorithm	96.2
Ref [6]	Voting	97.6
Ref [6]	Ensemble-based self-labeled algorithm (iCST-Voting)	98.4
Ref [17]	Multilayer perceptron deep learning model	96.8
Ref [25]	Logistic Regression	97.7
Ref [25]	kNN	97.7
Ref [25]	Naïve Bayes	89.4
Ref [25]	Decision Tree	96.7
Ref [25]	Random Forest	97.6
Ref [25]	SVM (linear kernel)	97.9
Ref [25]	Artificial Neural Network	98.4
Ref [25]	SVM (RBF kernel)	98.6
Ref [25]	Artificial Neural Network (keras based deep learning model)	99.8
This Method	Deeper LSTM (average test accuracy)	98.4
This Method	Deeper LSTM (maximum test accuracy)	100

When the studies in the literature using the same data are examined, it is seen that the study results are given only for the accuracy parameter. In addition, it is not stated whether the obtained accuracy values are the maximum result or the mean value. The proposed method has reached 100% accuracy in maximum results. It has an average value of 98.4%.

5. Conclusion and future works

In this study, an effective deeper LSTM networks structure was used by using a voice dataset. Before the classification process, 20 weights were calculated with Relieff method and the most effective 10 features were used. The classification was successful at 98.4% accuracy. Sensitivity and specificity values were also calculated. For the proposed method, these values were 97.2% and 99.5%, respectively. Traditional machine learning classifiers have also been studied to compare the performance of the proposed method. These classifiers are Tree, Linear Discriminant, Logistic Regression, SVM (with different activation functions) and KNN. Sensitivity and specificity values were calculated for each classifier. The best value for Sensitivity was obtained with KNN and Linear Discriminant with 97.4%. Sensitivity value was found as 97.2% with the proposed method. The classifier with the highest Specificity value is the proposed deeper LSTM classification and its value is 99.5%. The nearest classifier with this value was the tree classifier with 97.9%. In addition to accuracy, sensitivity and specificity performance metrics, geometric mean performance metric is calculated in this study. The highest geometric mean performance metric was 98.3%.

In future works, the existing algorithm will be tested by working with larger data sets. At the same time, it is aimed to achieve higher performance values by using a hybrid network with Bidirectional LSTM structure instead of LSTM. In addition, a web-based framework will be created. In the backend, a deep learning-based model will be run. This will enable users to upload audio data and make gender recognition.

References

[1] Li M, Jung C-S, Han KJ. 2010. Combining Five Acoustic Level Modeling Methods for Automatic Speaker Age and Gender Recognition. In: 11Th Annu. Conf. Int. Speech Commun. Assoc. 2010 (Interspeech 2010), Vols 3 4, 2010, p. 2830+.
[2] Chen C, Wang W, He Y, Han J. A bilevel framework for joint optimization of session compensation and classification for speaker identification. Digit Signal Process A Rev J 2019;89:104–15. https://doi.org/10.1016/j.dsp.2019.03.008.
[3] Qawaqneh Z, Mallouh AA, Barkana BD. Deep neural network framework and transformed MFCCs for speaker's age and gender classification. Knowledge-Based Syst 2017. https://doi.org/10.1016/j.knosys.2016.10.008.
[4] Black M, Katsamanis A, Lee C, Lammert AC, Baucom BR, Christensen A, et al. Automatic Classification of Married Couples' Behavior using Audio Features. Corpus 2010:2030–3.
[5] Nguyen P, Tran D, Huang X, Sharma D. Automatic speech-based classification of gender, age and accent. Lect Notes Comput Sci (including Subser Lect Notes Artif Intell Lect Notes Bioinformatics) 2010. https://doi.org/10.1007/978-3-642-15037-1_24.
[6] Livieris I, Pintelas E, Pintelas P. Gender Recognition by Voice using an Improved Self-Labeled Algorithm. Mach Learn Knowl Extr 2019;1:492–503. https://doi.org/10.3390/make1010030.
[7] Metze F, Ajmera J, Englert R, Bub U, Burkhardt F, Stegmann J, et al. 2007. Comparison of four approaches to age and gender recognition for telephone applications. In: ICASSP, IEEE Int. Conf. Acoust. Speech Signal Process. - Proc., vol. 4, 2007. doi:10.1109/ICASSP.2007.367263.
[8] Konig Y, Morgan N. 2003. GDNN: a gender-dependent neural network for continuous speech recognition, pp. 332–7. doi:10.1109/ijcnn.1992.226966.
[9] Acero A. Xuedong Huang. Speaker and gender normalization for continuous-density hidden Markov models 2002:342–5. https://doi.org/10.1109/icassp.1996.541102.
[10] Neti C, Roukos S. Phone-context specific gender-dependent acoustic-models for continuous speech recognition 2002:192–8. https://doi.org/10.1109/asru.1997.659005.
[11] Ramadhan MM, Sitanggang IS, Nasution FR, Ghifari A. Parameter Tuning in Random Forest Based on Grid Search Method for Gender Classification Based on Voice Frequency. DEStech Trans Comput Sci Eng 2017. https://doi.org/10.12783/dtcse/cece2017/14611.
[12] Harb H, Chen L. 2003. Gender identification using a general audio classifier. In: Proc. - IEEE Int. Conf. Multimed. Expo, vol. 2, 2003. pp. II733–6. doi:10.1109/ICME.2003.1221721.
[13] Huestegge SM, Raettig T. Crossing Gender Borders: Bidirectional Dynamic Interaction Between Face-Based and Voice-Based Gender Categorization. J Voice 2018. https://doi.org/10.1016/j.jvoice.2018.09.020.
[14] Zourmand A, Ting HN, Mirhassani SM. Gender classification in children based on speech characteristics: using fundamental and formant frequencies of Malay vowels. J Voice 2013;27:201–9. https://doi.org/10.1016/j.jvoice.2012.12.006.
[15] Barkana BD, Zhou J. A new pitch-range based feature set for a speaker's age and gender classification. Appl Acoust 2015;98:52–61. https://doi.org/10.1016/j.apacoust.2015.04.013.
[16] Pahwa A, Aggarwal G. Speech Feature Extraction for Gender Recognition. Int J Image, Graph Signal Process 2015;8:17–25. https://doi.org/10.5815/ijigsp.2015.09.03.
[17] Buyukyilmaz M, Cibikdiken AO. Voice Gender Recognition Using Deep Learning 2017. https://doi.org/10.2991/msota-16.2016.90.
[18] Zvarevashe K, Olugbara OO. 2018. Gender Voice Recognition Using Random Forest Recursive Feature Elimination with Gradient Boosting Machines. In: 2018 Int. Conf. Adv. Big Data, Comput. Data Commun. Syst. icABCD 2018, 2018. doi:10.1109/ICABCD.2018.8465466.
[19] Ay Karakuş B, Talo M, Hallaç İR, Aydin G. Evaluating deep learning models for sentiment classification. Concurr Comput 2018;30. https://doi.org/10.1002/cpe.4783.
[20] Hochreiter S, Schmidhuber J. Long Short-Term Memory. Neural Comput 1997;9:1735–80. https://doi.org/10.1162/neco.1997.9.8.1735.
[21] Lei J, Liu C, Jiang D. Fault diagnosis of wind turbine based on Long Short-term memory networks. Renew Energy 2019;133:422–32. https://doi.org/10.1016/j.renene.2018.10.031.
[22] Kong F, Li J, Lv Z. Construction of intelligent traffic information recommendation system based on long short-term memory. J Comput Sci 2018;26:78–86. https://doi.org/10.1016/j.jocs.2018.03.010.
[23] Bao W, Yue J, Rao Y. A deep learning framework for financial time series using stacked autoencoders and long-short term memory. PLoS ONE 2017;12:. https://doi.org/10.1371/journal.pone.0180944e0180944.
[24] Voice Dataset n.d. https://www.kaggle.com/primaryobjects/voicegender.
[25] Pondhu LN, Kummari G. Performance Analysis of Machine Learning Algorithms for Gender Classification. Proc Int Conf Inven Commun Comput Technol ICICCT 2018 2018. https://doi.org/10.1109/ICICCT.2018.8473192.

[26] Zhang H, Sun G. Feature selection using tabu search method. Pattern Recognit 2002;35:701–11. https://doi.org/10.1016/S0031-3203(01)00046-2.
[27] Kira K, Rendell LA. A Practical Approach to Feature Selection. Mach Learn Proc 1992;2014:249–56. https://doi.org/10.1016/b978-1-55860-247-2.50037-1.
[28] Jin L, Zeng Q, He J, Feng Y, Zhou S, Wu Y. A ReliefF-SVM-based method for marking dopamine-based disease characteristics: a study on SWEDD and Parkinson's disease. Behav Brain Res 2019;356:400–7. https://doi.org/10.1016/j.bbr.2018.09.003.
[29] Aviad B, Roy G. Classification by clustering decision tree-like classifier based on adjusted clusters. Expert Syst Appl 2011;38:8220–8. https://doi.org/10.1016/j.eswa.2011.01.001.
[30] Siqueira LFS. Araújo Júnior RF, de Araújo AA, Morais CLM, Lima KMG. LDA vs. QDA for FT-MIR prostate cancer tissue classification. Chemom Intell Lab Syst 2017;162:123–9. https://doi.org/10.1016/j.chemolab.2017.01.021.
[31] Schein AI, Ungar LH. Active learning for logistic regression: An evaluation. Mach Learn 2007;68:235–65. https://doi.org/10.1007/s10994-007-5019-5.
[32] Zhang S, Li X, Zong M, Zhu X, Wang R. Efficient kNN classification with different numbers of nearest neighbors. IEEE Trans Neural Networks Learn Syst 2018;29:1774–85. https://doi.org/10.1109/TNNLS.2017.2673241.
[33] Koopman B, Zuccon G, Nguyen A, Bergheim A, Grayson N. Extracting cancer mortality statistics from death certificates: a hybrid machine learning and rule-based approach for common and rare cancers. Artif Intell Med 2018;89:1–9. https://doi.org/10.1016/j.artmed.2018.04.011.
[34] Raghu S, Sriraam N. Classification of focal and non-focal EEG signals using neighborhood component analysis and machine learning algorithms. Expert Syst Appl 2018;113:18–32. https://doi.org/10.1016/j.eswa.2018.06.031.