Materials Today: Proceedings xxx (xxxx) xxx

Contents lists available at ScienceDirect

Materials Today: Proceedings

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

Feature based classification of voice based biometric data through Machine learning algorithm

Samiya Shakil [a,*], Deepak Arora [a], Taskeen Zaidi [b]

[a] Amity University, India
[b] Shri Ramswaroop Memorial University, India

ARTICLE INFO

Article history:
Received 30 March 2021
Received in revised form 3 May 2021
Accepted 11 May 2021
Available online xxxx

Keywords:
Prediction
Biometric data
Voice samples
Male
Female
Cost complexity pruning (CCP)
Dimension reduction

ABSTRACT

In the era of big data and growing artificial intelligence, the requirement and necessity of biometric identification increase in a rapid manner. The digitalization and recent Pandemic crisis gives a boost to need to authorized identification which get fulfilled with biometric identification. Our paper focuses on same concept of checking the identification accuracy of machine learning algorithm REPTree on selected biometric dataset which is being deployed and evaluated on a data mining tool WEKA. Our target is to achieve more or equal to 95 percentages in order to predict the given sample data is accurately classified into our target variables values i.e. male female. The selected algorithm REPTree is a kind of decision tree classification algorithm which works on same concept as C4.5 and decision tree algorithm with speciality of generation of both kind of output i.e. discrete and continuous. The selection of algorithm gives us benefits with achievement of higher accuracy and selection of dataset also become easy with some required modification and pre-processing of data with some dimension reduction filters.

Selection and peer-review under responsibility of the scientific committee of the 1st International Conference on Computations in Materials and Applied Engineering – 2021.

1. Introduction

Biometric authentication is the method of establishing a person's identity. Biometrics is made up of two words: First word is the Bio. This word menace the Greek word for life. Second word is the Metrics. The Metrics menace is the measurement of something. This word is the Measurements. It is the branch of information technology aimed at defining an individual's individuality founded on personal characteristics biometric of image recognition has been used where prediction and feature extractions are the main methods which have been applied by using data mining algorithms [1].

Biometrics is a technique used to recognize, analyzing, and measurement an individual's behavioral and physical characteristics.

- The aim is to find biometric data from this individual. It may be a photograph of their face, a voice recording, or a fingerprint image.
- This data is then related to biometric data from a large number of another people stored in a database.

The process of comparing an individual's unique biometric parameter to the entire database of available data is known as biometric recognition. This type of personal data is commonly retrieved using biometric readers [3].

There are two types of the biometrics which is available:

- First biometric is the physiological.
- Second biometric is the Behavioural.

2. Physiological measurements

It is the first type of the biometric. They may be biological or morphological in nature. Physical recognition approaches are based on a thorough examination of a person's invariable physiological characteristics.

- The shape of the palm, Fingerprints, vein pattern, the finger, the eye (retina and iris), and the shape of the face are the most common morphological markers.

* Corresponding author.
E-mail address: samiyashakil316@gmail.com (S. Shakil).

https://doi.org/10.1016/j.matpr.2021.05.261
2214-7853/
Selection and peer-review under responsibility of the scientific committee of the 1st International Conference on Computations in Materials and Applied Engineering – 2021.

- Medical teams and police forensics can usage blood, DNA, urine, or saliva for biological analyses.

1. Behavioral identification methods

The study of an individual's behavioral characteristics — the characteristics inherent in each person in the course of reproducing an action — is the foundation of behavioral recognition methods [2].

Voice Recognition or Speaker Recognition:

One of the types of biometric authentication is voice recognition, which allows an individual to be identified using a grouping of specific voice attributes and relates to dynamic biometric approaches. Voice recognition is a technique for mechanically identifying a speaker depended on the speech waveform, which represents the behavioural and physiological features of the speaker's speech factors [4].

There are two phases to this method, similar to conventional speaker recognition systems: preparation and testing. These are the key phases in the identification of a speaker. Learning is the method of extracting phonetic attributes from a registered or stored sample of a speaker, storing them in a database, and familiarizing the machine with the speaker's voice characteristics. The method of matching suspect sound and phonetic characteristics from a voice recognition database is known as testing [5].

Verification and identification are two types of speaker recognition. There are two levels to the Speaker Recognition scheme. Verification and recognition of the speaker. The voice print of a speaker is paired with a single prototype in a 1:1 match. Speaker recognition, on the other hand, is a 1:N match, in which the input speech is matched against multiple models. There are five types of the speaker verification which is below:

1. First speaker verification is the Input data acquisition.
2. Second speaker is the feature extraction.
3. Third speaker is the pattern matching.
4. Forth speaker is the decision making.
5. Fifth speaker is the generate speaker models.

In the first stage, the user's speech is sampled in a controlled manner. The speech signals will be processed by the speaker recognition system, which will extract speaker discriminatory information. This data is used to build a speaker model. A sample voice print is obtained from the user during the verification process. The characteristics of the input speech will be extracted by the speaker recognition system. This procedure is termed as the Pattern matching [6].

Speaker recognition's main goal is to transform an acoustic audio signal into a computer-reliable format. Training and testing are two phases of speaker recognition systems. Take the input as a speech signal during the training phase, and use the feature extraction technique to extract features. Feature vectors represent the speaker's voice characteristics and are used to construct a reference model [1]. The actual recognition task is still being tested. During the testing process, a matching technique is used to adapt the speaker voice to the reference model. After the decision on the degree of matching is done [7].

REPTree:

RepTree builds several trees in various iterations using regression tree logic. After that, it chooses the best tree out of all the ones that have been created. This will be taken into account as the representative. The mean square error on the tree's predictions is the metric used to prune the tree. "REPT" is a quick decision tree learning method that constructs a decision tree depended on knowledge gain or variance reduction. Here, REPT is referred as the Reduced Error Pruning Tree. This method is a quick decision tree learner that splits and prunes a regression/ decision tree using knowledge gain as the splitting measure. This only sorts numeric attribute values once [6].

"REPT" is a quick decision tree learning method that constructs a decision tree based on knowledge gain or variance reduction. Here, REPT is referred as the Reduced Error Pruning Tree. The simple principle of this algorithm's pruning is that it employs Reduced Error Pruning Tree with back over fitting [8].

Many of the branches in a decision tree will replicate deviations in the training data due because of outliers or noise. The problem of overfitting the data is addressed by tree pruning methods. Here, There are 2 types of the traditional methods to tree pruning which is below:

1. First method is the pre-pruning.
2. Second method is the post-pruning.

These techniques usually use statistical tests to eliminate the least stable branches. Pruning's main motive [2] is to "trade precision for simplicity." Pruning decision trees can be done in a variety of ways. The majority of them traverse the nodes from the top down or bottom up. If this operation strengthens a certain criterion, a node is pruned [9].

Quinlan [1987b] proposes a method for producing a set of pruned trees through explicitly using test data instead of consuming it only for selecting the best tree.

This is how it's done: Begin by creating a full tree and running the test data through this, noting the quantities that appear in - class at each node. Count the number of errors for each non-leaf node when the subtree is held, and the number when this becomes a leaf using pruning method. On the test results, the pruned node will often make fewer errors than the sub-tree. The benefit from pruning the sub-tree is measured by the variance among the quantities of errors (when positive). Choose the node with the greatest difference as the sub-tree to prune from all of the nodes. Continue this method, including those nodes where the reduction is zero, until the rate of misclassification increases. With respect to the test set, this yields the smallest version of the most reliable tree [10].

It's possible that there are many sub-trees with the same (largest) variation. Quinlan [1987] does not say which sub-tree to select in this case, such as the largest or smallest. Experiments show that the option makes no difference in classification accuracy; hence, the largest is chosen because it decreases the number of iterations needed to fully prune the tree. This method produces a series of trees, with the smallest minimum-error tree on the test data being the final one (Fig. 1Fig. 2. Fig. 3. Fig. 4. Fig. 5. Fig. 6. Fig. 7. Fig. 8. Table 1)

A. Cost-Complexity Pruning

There are two stages to cost complexity pruning. Cost complexity also referred as the error complexity pruning or weakest connection pruning. Sequences of trees are based on the training datasets in the first step, with the root tree being the initial tree before pruning. One of these trees is selected as the pruned tree in the second level, based on its generality of error estimation.

$$\mathrm{ERT}(t) = \mathrm{n}(t) \ * \mathrm{s}(t)$$

Here n(t) is number of misclassified instance divided by all instances present in node

And p(t) is total number of instance in a node divided by all present node

B. Pessimistic Pruning

Instead of using a cross validation or pruning range, pessimistic pruning employs a pessimistic statistical association examination. The simple premise is that the error ratio calculated through the training set is insufficiently compatible. Instead, a more realistic measure for binomial allocation referred as "continuity adjustment" should be used [11].

C. Reduced-Error Pruning

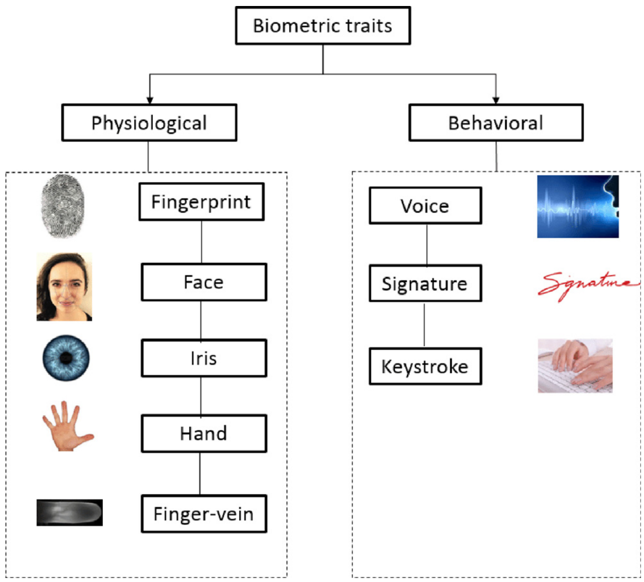


Fig. 1. Classification of biometric traits (Sample categorisation).

The REP [5] method is similar to traversing a tree's internal nodes from bottom to top. Checks if replacing every internal node with the most frequently occurring class reduces tree accuracy. The node is pruned in this situation. The process is repeated before any further pruning will result in a reduction in accuracy. Quinlan recommends using a pruning range to estimate the precision. This can be exposed that this method produces the minimum precise subtree for a specified pruning set [12].

3. Literature survey

According to article discusses the classification algorithms for the problem of voice-based personality recognition through MI approaches. Here, MI is referred as the Machine Learning. The Mel Frequency Cepstral Coefficient algorithm was used in the speech pre-processing technique. A comparative study of five classification algorithms was conducted to solve the problem. In the first experiment, the best results were obtained using the support vector method (0.90) and the multilayer perceptron (0.83). In the second experiment, the Robust scaler approach was used to propose a multilayer perceptron with an accuracy of 0.93 for personal recognition [13].

According to R. Shiva Shankar, J. Raghaveni, Pravallika Rudraraju, Y. Vineela Sravya, R. Shiva Shankar, "Voice Recognition for Gender Classification Using Machine Learning Algorithms" The gender of an individual has become increasingly important in the economic markets, especially in the form of advertising. The aim of this project is to create a device that can determine the gender of a speaker based on the pitch of their speech. Machine learning can be used to identify the gender from the properties of a voice data set, such as pitch, median, frequency, and so on. In this project, we're attempting to identify gender as male or female based on a dataset containing various voice attributes such as pitch, frequency, and so on. To find the gender classification of voice data using machine learning algorithms, data pre-processing steps should be performed [14].

Emmanuel Pintelas 1 and Panagiotis Pintelas 2 are Ioannis E. Livieris 1, Emmanuel Pintelas 1 and Panagiotis Pintelas 2 respectively. Speech Gender Recognition Using a Self-Labeled Algorithm that has been Improved Speech recognition has a wide range of applications, including human–machine interaction, gender categorization of phone calls, video categorization with marking, and so on. MI is currently a mainstream trend that has been widely used in a variety of fields and applications, taking advantage of recent advances in digital technology and the advantages of electronic media's storage capabilities. Here MI is referred as the Machine Learning. To produce more precise classifiers, researchers have recently focused on using ensemble learning methods in conjunction with a semi-supervised learning system. In this paper, we tackle gender recognition by speech using a new ensemble semi-supervised self-labelled algorithm. The proposed algorithm's classification efficiency in terms of accuracy is demonstrated in our preliminary numerical experiments, paving the way for the development of reliable and robust predictive models [15].

3.1. Implementation and experimental evaluation

For implementation and calculation of accuracy score in order to correctly identify our target variable i.e. gender we had break the complete experiment into several steps, which are as follows:

1. Preparation and pre-processing step of dataset
2. Selection of machine learning algorithm
3. Apply cross validation and selection of number of folds to be applied on prepared dataset

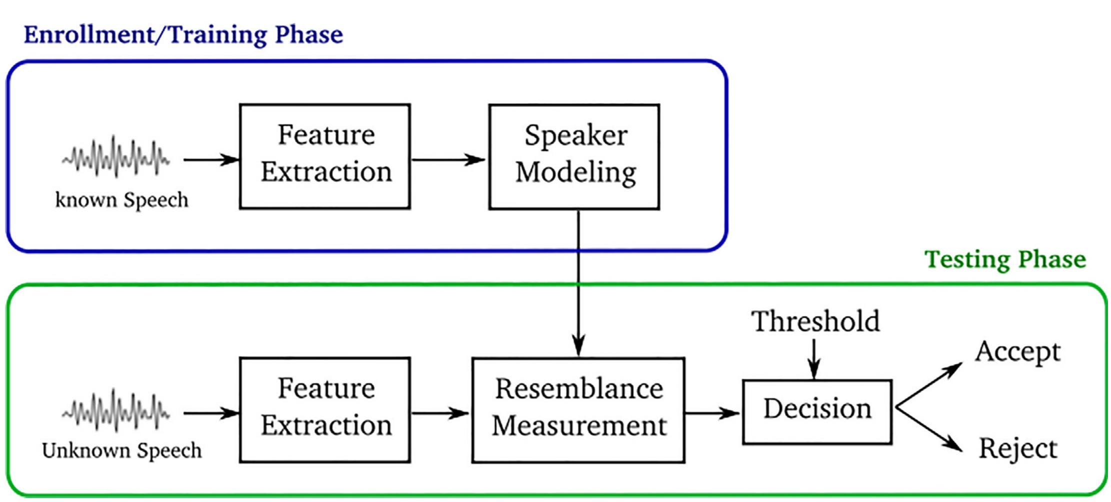


Fig. 2. The process of training phase and Testing phase.

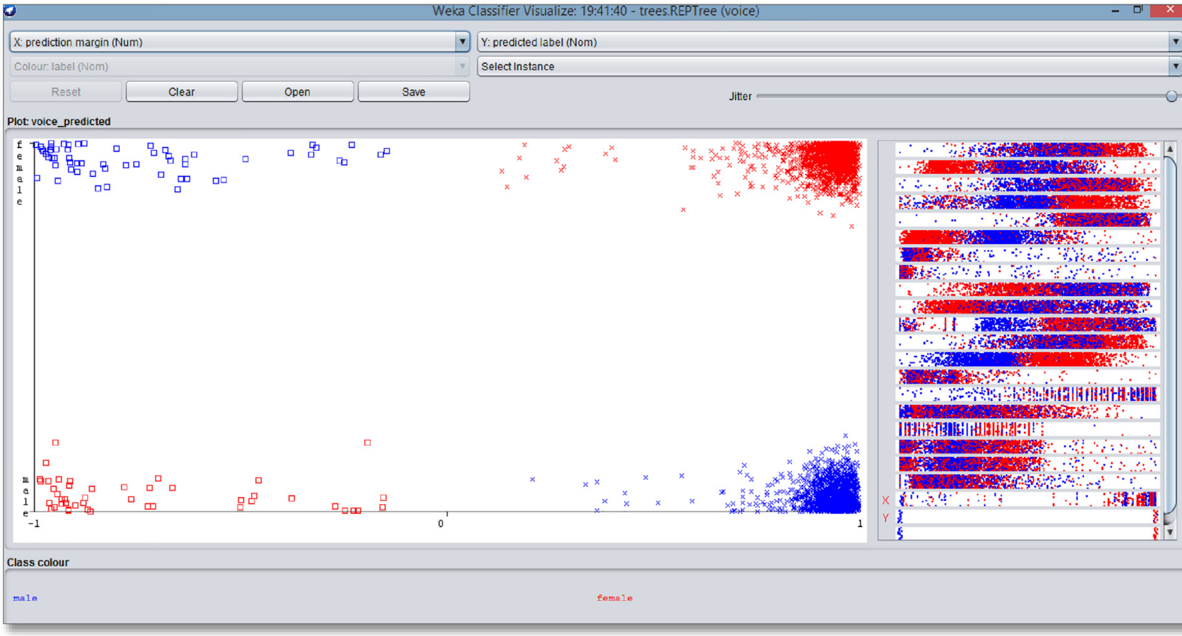


Fig. 3. Visualization of classification of our class variable gender (values: male, female).

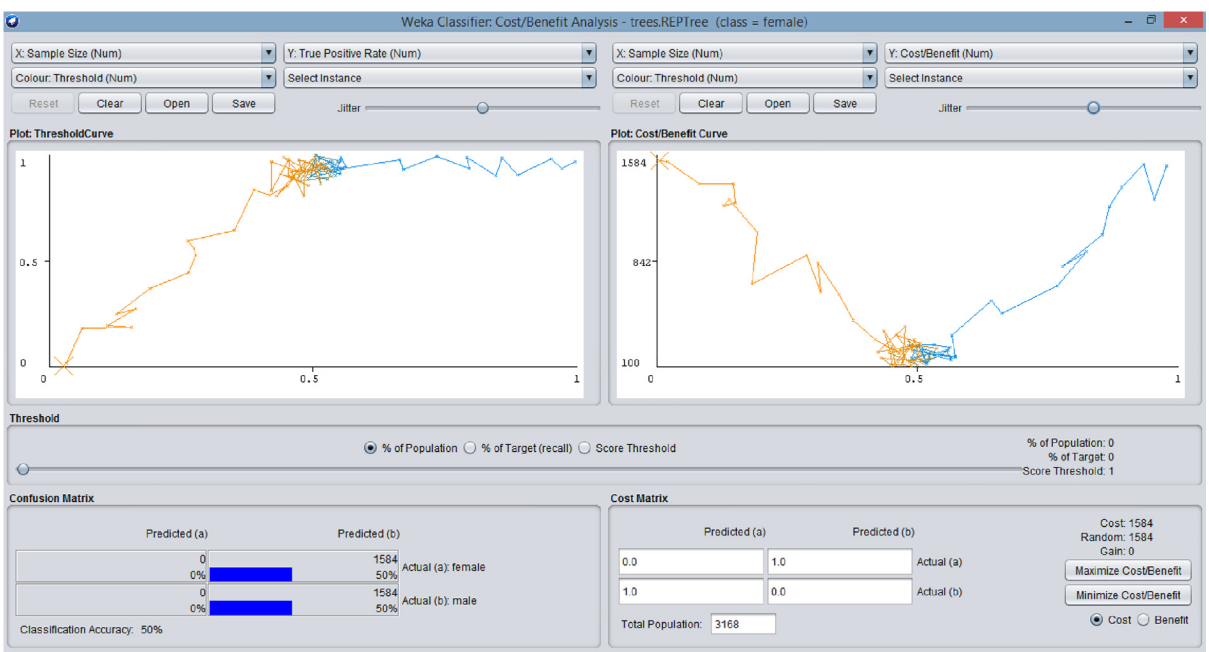


Fig. 4. Weka classifier cost/benefit analysis for target label female.

4. Evaluation of results
5. Preparation of accuracy parameter table
6. Aggregating all the output and threshold values
7. Final conclusion

Preparation and pre-processing steps of dataset:

Pre-processing and preparation of dataset is the most basic and initial stage while we want to prepares a model of classification or prediction purpose. Here the selection of attribute and and the number of instance which we can considering in our final dataset is really important, also several other factors are also important to save the model from un-necessary overhead of curse of dimensionality which is the overhead on model due to excessive and unnecessary dimension present in training dataset [16].

Here we have applied two dimension reduction filters to clean our data from unwanted dimensions. We have removed the blank and null values from dataset and then we check if some attribute are having low correlation with the target variable then we also removed that. Finally the selected instance in dataset is 3168 with refined 21 attributes.

Selection of machine learning algorithm:

The next step after pre-processing and dimension reduction of dataset, we have a dataset which is ready of preparing and training a model. Now next step is to choose a suitable machine learning

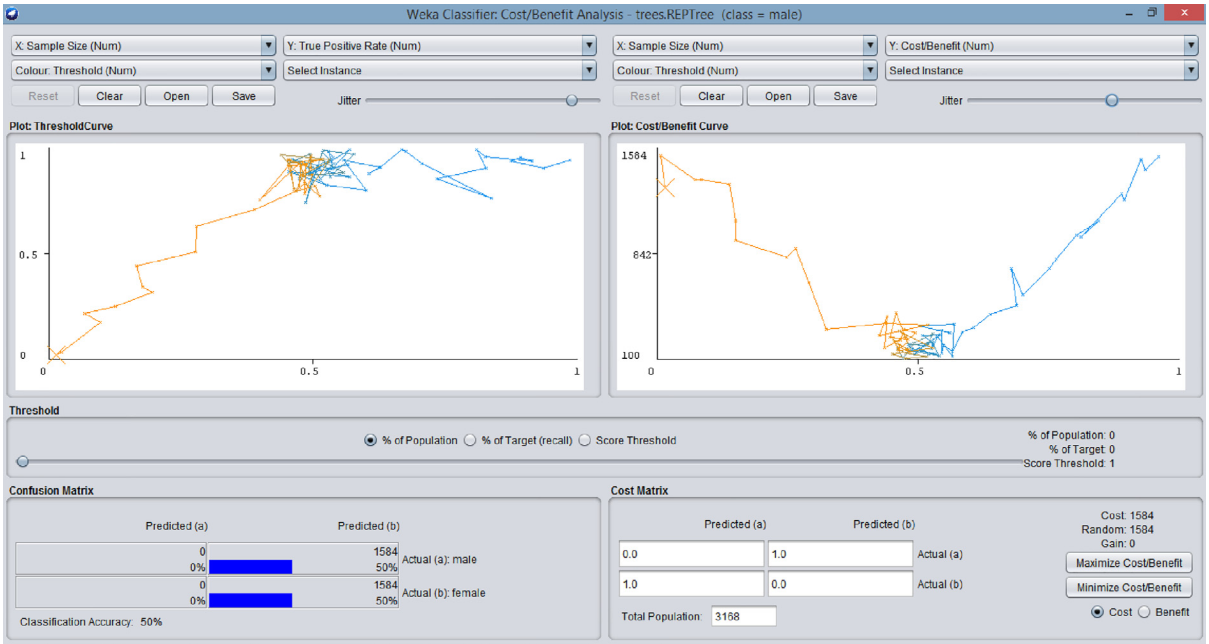


Fig. 5. Weka classifier cost/benefit analysis for target label female.

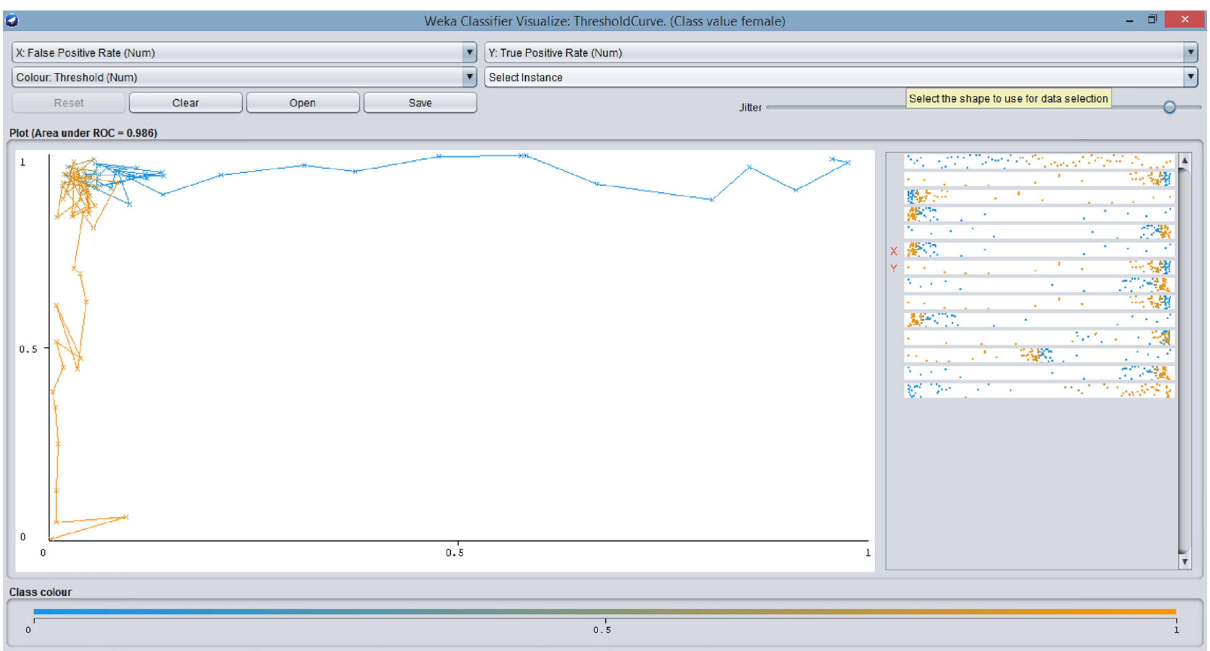


Fig. 6. Weka classifier ThersholdCurve for class value female.

algorithm which is giving best accuracy score for our dataset, for our paper we have chosen REPTREE algorithm for prediction of gender on the basis of voice sample dataset. There are several advantages of using this algorithm including additional features of C4.5 algorithm and decision algorithm with ability to process all type of data like continuous and discrete data [16].

Apply cross validation and selection of number of folds to be applied on prepared dataset:

We have applied cross validation while training the dataset model, as the cross validation is process to achieve the accessibility to each dimension and view of dataset which is not being seen in normal train test splitter with criteria of 70 to 30 ratio. Here the cross validation allow us to view the insight of dataset from several view and we can set the number of folds with k value which can range from 1 to 100 and more if you want. In our experiment we have set k value to 10 which is optimal for our model [17].

4. Evaluation of results:

5. Preparation of accuracy parameter table

6. Checking all evaluation parameters with cost benefit analysis and threshold values.

7. Final REPtree:

The final output of REPTree can also be visualized as tree hierarchy; the output of REPTree has size of 33 nodes. The selection of root node is done with criteria of attribute of being able to split

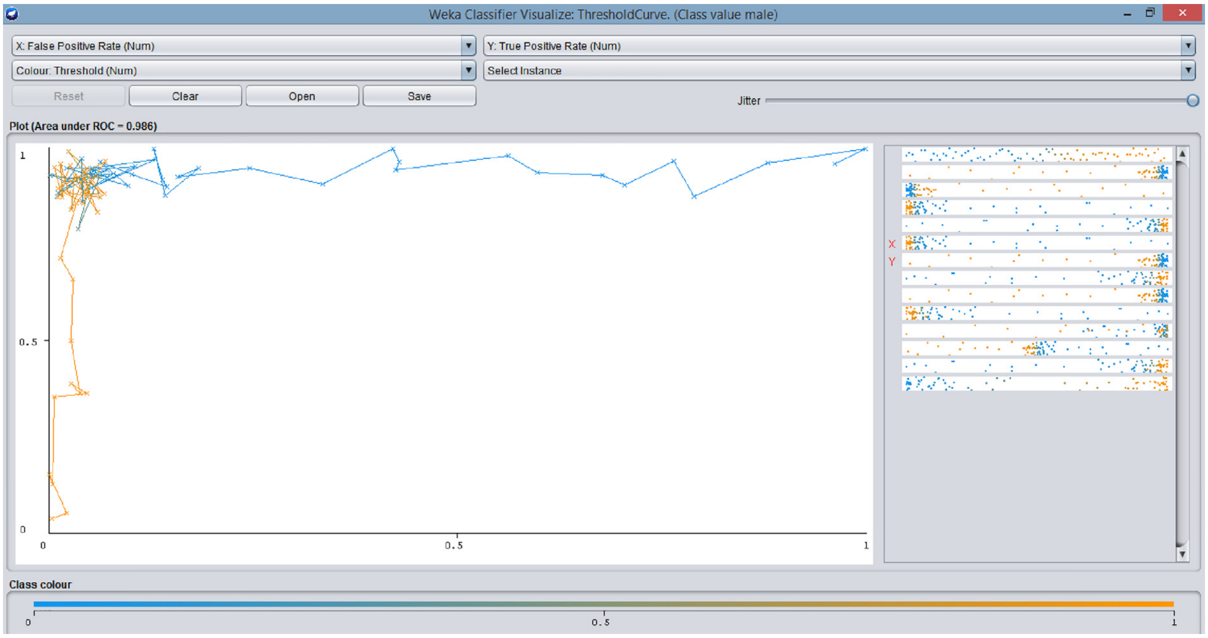


Fig. 7. Weka classifier ThersholdCurve for class value male.

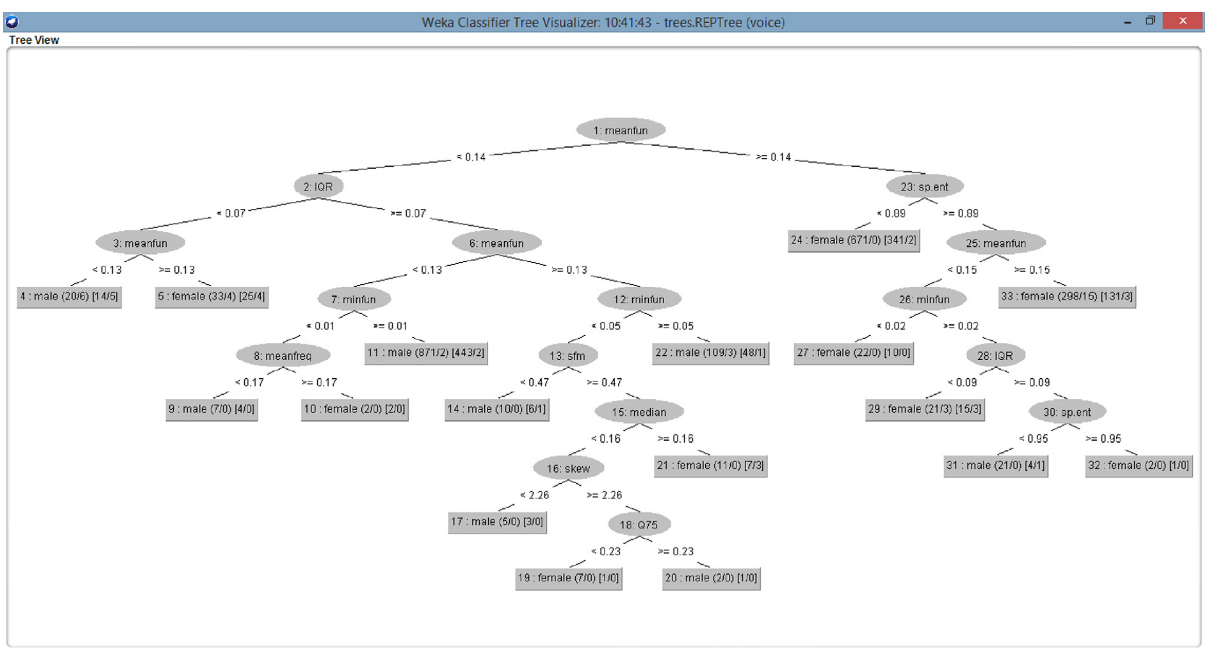


Fig. 8. Final hierarchy after implementation of REP Tree (Tree size: 33).

Table 1
Detailed accuracy measure after implementation of REPTree algorithm.

TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC area	PRC Area	Class
0.963	0.029	0.971	0.963	0.967	0.934	0.986	0.984	Male
0.971	0.037	0.963	0.971	0.967	0.934	0.986	0.981	Female
0.967	0.033	0.967	0.967	0.967	0.934	0.986	0.983	Weighted Avg.

the data for other attribute to achieve target value i.e. male or female. Here mean fun is selected as root node under splitting criteria of 0.14 ad further splitting is done until the attribute reach to our target value either male or female.

Output: For identification of biometric data i.e. voice sample recognition using machine learning algorithms we have choose REPTree algorithm for our prediction purpose. The choice of algorithm means with selected algorithm we have achieved 96.68%

accuracy score on a dataset of 3168 instances with acceptable level incorrectly classified instance of 3.31%. Hence the implementation of algorithm was a right choice for our voice sample dataset and accuracy score of prediction is clearly high, it was a successful implementation [17]. Various factors of measurement of algorithm accuracy like time taken in model building was minimal i.e 0.16 s and for both target labels male and female the true positive rate is higher as expected and the false positive rate is minimal which boost the confidence of result accuracy [18]. With this existing algorithm module various other biometric data skin retina sample dataset, thump impression dataset and many kind of dataset can also be evaluated in future for having more accuracy score. And also there is a scope of enhancement in the dataset instances which will surely give best results in prediction [19].

4. Future work

[A 163-page report] From USD 7.5 billion in 2018 to USD 21.5 billion in 2024, the voice and speech recognition market is expected to grow at a CAGR of 19.18 percent. The market's development is fuelled by aspects For instance rising demand for voice authentication in mobile banking apps, the rapid proliferation of smart speakers or multifunctional devices, and the rising effect of artificial intelligence on voice and speech recognition accuracy [18].

Speaker recognition is expected to expand at a faster pace than speech recognition from 2018 to 2024, owing to the increasing usage of Speaker recognition in multifactor authentication systems in the BFSI, government, and security verticals. A significant number of banking customers in North America and Western Europe use phone banking services [19]. To accept or deny a user's mobile transactions, many of these financial institutions are implementing voice-based authentication solutions. Furthermore, over the next 2–3 years, the demand for voice recognition technology is expected to grow rapidly in the government, banking, and business verticals. Concerns about data protection as a result of cyberattacks, data breaches by intruders, and issues relating to illegal migrants are just a few of the main aspects driving the voice recognition market's rapid development.

CRediT authorship contribution statement

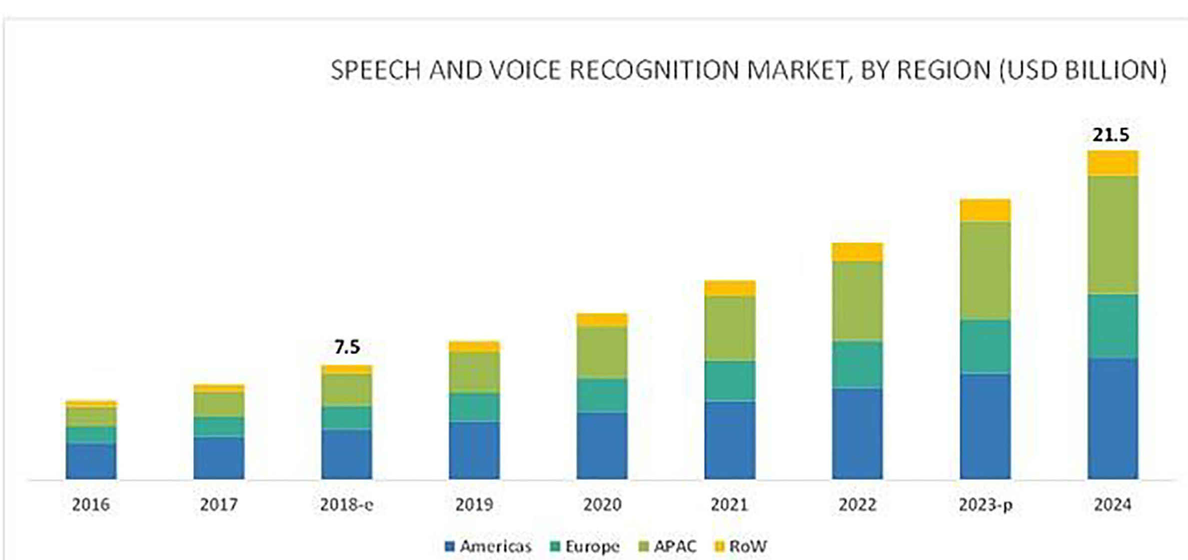
Samiya Shakil: Investigation, Writing - original draft. **Deepak Arora:** Conceptualization, Supervision, Writing - review & editing. **Taskeen Zaidi:** Data curation, Formal analysis.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References:

[1] According to K.A. Senthildevi, E. Chandra. Data mining techniques and applications in speech processing-A Review, IJARS 1 2012 1-8.
[2] According to V. Bharat, B. Shelale, K. Khandelwal, S. Navsare. A review paper on data mining techniques, Int. J. Eng. Sci. Comput. 6 2016 6268-6271.
[3] According to M. Buyukyilmaz, A.O. Cibikdiken. Voice gender recognition using deep learning, Adv. Comput. Sci. Res. 58 2016 409-411.
[4] According to A. Jha, M. Dave, S. Madan. A reviews on the study and analysis of big data using data mining techniques, IJLTET 6 2016 94-102.
[5] S.V. Manikanthan, K. Baskaran. Low Cost VLSI Design Implementation of Sorting Network for ACSFD in Wireless Sensor Network, CiiT Int. J. Programmable Device Circuits Syst. Print: ISSN 0974 – 973X & Online: ISSN 0974 – 9624, Issue : November 2011, PDCS112011008.
[6] According to J. Pribil, A. Pribilova, J. Matousek. GMM-based speaker gender and age classification after voice conversion. In First international workshop on sensing, processing and learning for intelligent machines (SPLINE), Denmark. 2016.
[7] M. Fezari, F.Z. Taif, M.M. Lafifi, B. Boulebtateche, Noise emission analysis a way for early detection and classification faults in rotating machines, in Power Electronics and Motion Control Conference and Exposition (PEMC), 2014 16th International, 2014, pp. 1094-1099.
[8] Qiang Huo, Chin-Hui Lee, A bayesian predictive classification approach to robust speech recognition, IEEE Trans. Speech Audio Process. 8 (2) (2000) 200–204.
[9] J.R. Bellegarda, Large vocabulary speech recognition with multispan statistical language models, 76 IEEE Trans. Speech Audio Process. 8 (1) (2000) 76–84.
[10] C.H. Lee. Statistical confidence measures and their applications, Proc. ICSP 2001 1021-1028.
[11] G. Saon, M. Padmanabhan, Data-driven approach to designing compound words for continuous speech recognition, IEEE Trans. Speech Audio Process. 9 (4) (2001) 327–332.
[12] According to T. Padmapriya, V. Saminadan. Inter-cell Load Balancing technique for multi-class traffic in MIMO-LTE-A Networks Int. J. Electric. Electron. Data Commun. (IJEEDC), ISSN: 2320- 2084, 3(8) 2015 22-26.



[13] Mohamed Afify, Feng Liu, Hui Jiang, A new verification-based fast-match for large vocabulary continuous speech recognition, IEEE Trans. Speech Audio Process. 13(4) 2005.
[14] F. Wessel, H. Ney, Unsupervised training of acoustic models for large vocabulary continuous speech recognition, IEEE Trans. Speech Audio Process. 13 (1) (2005) 23–31.
[15] According to H. Sakoe, Two Level DP Matching A Dynamic Programming Based Pattern Matching Algorithm for Connected Word Recognition , IEEE Trans. Acoustics, Speech, Signal Proc., ASSP-27:588-595, December 1979.
[16] L. Rabiner, S. Levinson, A. Rosenberg, J. Wilpon, Speaker independent recognition of isolated words using clustering techniques, IEEE Trans. Acoust. Speech Signal Proc. 27 (4) (1979) 336–349.
[17] C.C. Tappert, N.R. Dixon, A.S. Rabinowitz, and W.D. Chapman, Automatic Recognition of Continuous Speech Utilizing Dynamic Segmentation, Dual Classification, Sequential Decoding and Error Recover , Rome Air Dev.Cen, Rome, NY,Tech.Report TR-71- 146,1971.
[18] D.B. Fry, Theoritical Aspects of Mechanical speech Recognition, P. Denes, The design and operation of the mechanical speech recognizer at Universtiy College London, J. Br. Inst. Radio Eng. 19(4) 1959 211-299.
[19] According to Sadaoki Furui, 50 years of Progress in speech and speaker recognition research, ECTI Trans. Comput. Inf. Technol. 1(2) 2005.