Materials Today: Proceedings xxx (xxxx) xxx

Contents lists available at ScienceDirect

Materials Today: Proceedings

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

Voice recognition system using machine learning techniques

Ashraf Tahseen Ali *, Hasanen S. Abdullah, Mohammad N. Fadhil

Department of Computer Science, University of Technology, Baghdad, Iraq

ARTICLE INFO

Article history:
Received 30 March 2021
Accepted 5 April 2021
Available online xxxx

Keywords:
Machine learning
Voice recognition
Naïve Bayes
k-Nearest Neighbor
MFCC

ABSTRACT

Voice is a Special metric that, in addition to being natural to users, offers similar, if not higher, levels of security when compared to some traditional biometrics systems. The aim of this paper is to detect impostors using various machine learning techniques to see which combination works best for speaker recognition and classification. We present several methods of audio preprocessing, such as noise reduction and vocal enhancements, to improve the audios available in real environments. Mel Frequency Cepstral Coefficients (MFCC) are extracted for each audio, along with their differentials and accelerations, to verify machine learning classification methods such as PART, JRip, Nave Bayes, RT, J48, Random Forest, and k-Nearest Neighbor Classifiers. examine the 7 classifiers on two datasets, the extent of accuracy achieved for each classifier. Among the high performance were the random forest algorithm and the naive bias algorithm, and the weak performance of the PART algorithm.

Selection and peer-review under responsibility of the scientific committee of the Emerging Trends in Materials Science, Technology and Engineering.

1. Introduction

Finger Voice can combine what people say and how they say it by two-factor authentication in a single action. Other forms of identification can help with biometrics, but voice identification is needed for safe and unique authentication. Personal voice recognition and telephone recognition are two variables that can be combined with voice [1]. Voice recognition systems are inexpensive and easy to use. In today's smart world, voice recognition is crucial in a variety of ways. Voice-activated banking, home automation, and voice-activated devices are only a few of the many uses for voice recognition [2]. The process of recognizing a person based on his voice signal is known as speaker recognition. Because of variations in the shape of the vocal tract, the size of the larynx, and other sections of the voice production organs, each person's sound may be unique [3]. Since voice recognition must be conducted in a variety of environments, the features extracted must also be resistant to background noise and sensor mismatches [4]. the speaker's voice to be used to verify their identity and monitor access to services like voice dialing, telephone banking, dataset access services, information service, voice mail, and security control for sensitive information fields, and remote device access [5].

2. Literature survey

For sixty years, research in automated speech recognition by machines has attracted a lot of interest for a variety of reasons ranging from scientific curiosity about the tools for the mechanical realization of human speech abilities to a request to automate manageable tasks that demand human–machine interactions [6]. In this section, some of the previous work related to this research will be reviewed:

In 2017, the researchers have proposed a recognition systems are implemented using both spectro-temporal features and voice-source features. For the i-vector process, classification is performed with two separate classifiers, and the accuracy rates are compared. It was decided to compare the efficiency of two separate speaker recognition systems. It is evident from the study that GMM performs better than i-vectors in the case of short utterances, with an accuracy of 94.33%, and that there was a substantial improvement in the accuracy rates when concatenated test signals were used [7]. In 2018, the researcher proposed speech recognition system using SVM. Individual words are separated from continuous speeches using Voice Activity Detection (VAD). Each isolated word's features were extracted, and the models were successfully educated. Each individual utterance is modelled using SVM. The MFCC is used to describe audio content and is measured as a collection of features. By learning from training data, the SVM learning algorithm was used to recognize speech. The proposed audio

* Corresponding author.
E-mail address: ashraf88ashraf888@gmail.com (A. Tahseen Ali).

https://doi.org/10.1016/j.matpr.2021.04.075
2214-7853/
Selection and peer-review under responsibility of the scientific committee of the Emerging Trends in Materials Science, Technology and Engineering.

support vector machine learning system has a strong output in 95% speech recognition score, according to experimental results [8]. In 2019, the researchers have suggested a system to recognition and identification in Arabic speaker. It is divided into two phases (training and testing), each of which involves the use of audio features (Mean, Standard Division, Zero Crossing, Amplitude). Following the feature extraction, the recognition stage employs (J48, KNN, LVQ), with the Nearest Neighbor (KNN) neural network used for data training and testing, and the LVQ neural network for Speech Recognition and Arabic Language Identification. They had a higher recognition rate of 85, 93, and 96.4% [9]. In 2020, the researchers aim to test various pre-processing, feature extraction, and machine learning techniques on audios captured in unconstrained and natural settings to see which combination of these works best for speaker recognition and classification. This work is divided into three sections: audio preprocessing, feature extraction (in which Mel Frequency Cepstral Coefficients (MFCC) are extracted for each audio), and machine learning classification (using the Random Forest Algorithm to obtain the best classification rate for its hyper-parameter). The accuracy of using RF classifier reached 84% [10].

3. The proposed system architecture

The proposed system which is using biometrics will be recognized voice depending on machine learning system. In general, consist of voice records, dataset description, pre-processing, feature extraction and classification stages and post-processing stage. The proposed system architecture is as shown in Fig. 1.

3.1. Database description

Input part is prerequisite for a voice recognition system. WAV and MP3 are the two most common audio formats currently available. WAV files are preferred by most researchers because they span the full spectrum of frequencies audible to the human ear. MP3 files, on the other hand, are compressed and hence do not contain all of the information that a WAV file of the corresponding audio does. Furthermore, function extraction from these WAV files is critical. This step serves as the foundation for the machine learning algorithms that will be used to classify the data. As a result, WAV files are often used in audio studies. In an audio sample, consistency in sampling rate is critical to ensure that the extracted coefficients reflect the same underlying calculations. In this work two data sets were used. The first voices dataset was (Prominent leader's speeches), Includes audio clips of five country leaders, the second called Speaker Recognition Audio Dataset, Contains audio clips of fifty persons. Both were downloaded and the details from the kaggle website, and the details of these data are shown in the Table 1.

3.2. Pre-processing

The preprocessing stage's main advantage is that it organizes the data, making the recognition task easier. All operations relating to audio are referred to as "preprocessing."

3.2.1. Remove noise using hamming window

Windowing is a method of analyzing long sound signals by selecting a sufficiently representative segment [11]. This process using to removes noise in a signal that is polluted by noise present in a wide frequency spectrum:

$$y_1(n) = x_1(n)\ w(n), 0 \leqslant n \leqslant N-1 \quad [11] \tag{1}$$

where

y(n): is the product of the convolution between the input signal and the window function.
x(n): is the signal to be convolved by the window function.
w(n): usually uses window hamming which has the form.

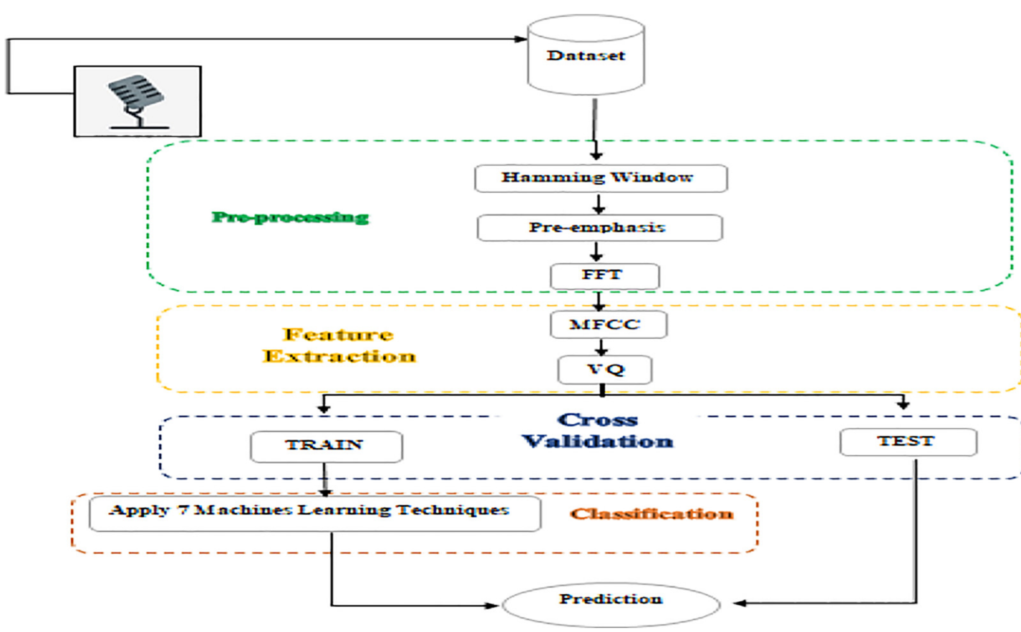

Fig. 1. The propose voice recognition system architecture.

Table 1
System's Datasets.

Input file Data	Name of Dataset	File format	File size	No. of sample
Voice Recognition	Prominent leaders s speeches	.Wave	16 khz	7500
	Speaker Recognition	.Wave	16 khz	2226

3.2.2. Smoothing spectral of speech signal using Pre-emphasis

The pre-emphasis filter is needed for speech signal processing. The pre-emphasis filter is based on the time domain input/output relationship expressed in the equation below [12]. The aim of using this filter is to make the spectral form of the speech signal frequency more smooth, this process apply by using the Eq. (2) blowe:

$$y(n) = x(n) - \ ax(n-1) \quad [12] \tag{2}$$

where, a is a pre-emphasis filter constant, it is usually 0.9 < a < 1.0.

3.2.3. Signal domain transform based on (FFT)

The Fourier series can be used to express a function with a finite time. A time series of bounded time-domain signals is converted into a frequency spectrum using the Fourier transform [13]. This process used to convert each frame from the time domain to the frequency domain. This process done using Eq. (3) below:

$$X_n = \sum_{k=0}^{N-1} x_k e^{-2\pi jkn/N} \quad [13] \tag{3}$$

where

x_k: is the signal of a frame
X[n]: is the n frequency pattern formed by the Fourier transform.

3.3. Feature extraction

Feature extraction is the process of calculating a collection of feature vectors that provides a compact representation of a particular speech signal.

3.3.1. Apply Mel-Frequency Cepstral Coefficients (MFCC)

MFCC is a method that uses human hearing activity to detect frequencies above 1 kHz. The MFCC system is focused on the frequency differences that the human ear can detect [14]. The number of Cepstral Coefficients was selected is 12 which results in more complexity in the voice proposed system.

3.3.2. Apply vector quantization

Quantization is an unavoidable phase in the digital representation of signals for computer processing [15]. Here, it was used to convert the binary matrix created by MFCC to a one-row matrix until it was combined with the other tools' output matrices (Fig. 2).

3.4. Proposed system classifiers

Machine learning classifiers, including feature extraction techniques, are critical in assessing the overall effectiveness of the speaker recognition model. This is a classification problem since we want to classify audios and figure out who is speaking in them. As a result, the following successful supervised classification machine learning algorithms will be used.

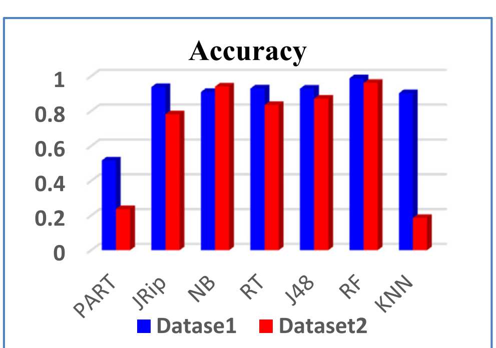

Fig. 2. Accuracy measured for Classifiers.

3.4.1. PART

PART is a separate-and-conquer rule learner. The algorithm generates "decision lists," which are pre-determined sets of rules [16]. This algorithm produces a decision list, which is an ordered set of rules. Each rule in the list is compared to new data, and the data is assigned to the category of the rule with the best match.

3.4.2. JRIP

One of the most common and widely used algorithms is JRip. Classes are analyzed as they grow larger, and an initial set of rules for the class is created using incrementally lower error rates [17]. This algorithm used to classify all of the examples of a given dataset in the training data and seeking a set of rules that apply to all members of that class. It then moves on to the next class and repeats the process until all classes have been examined.

3.4.3. Naïve Bayes (NB)

The Naïve Bayes classifier is a straightforward probabilistic classifier based on Bayes' Theorem and strict independence assumptions as shown in Eq. (4), assuming that all features are equally independent [18]. The feature will be assigned to the class of posterior probability using the NB classifier of the probability that the feature belongs to a class of prior probability. The consequence of prediction is the class with the highest posterior probability. This classifier predicts the test data set's class quickly and accurately, and it also performs well in multiclass prediction.

$$p(c|x) = \frac{p(x|c)p(c)}{p(x)} \quad [18] \tag{4}$$

where

$p(c|x)$: the posterior probability of class (c, target) given predictor (x, attributes).
$p(c)$: the prior probability of class.
$p(x|c)$: the likelihood which is the probability of predictor given class.
$p(x)$: the prior probability of predictor.

3.4.4. REP tree (RT)

The REPTree is an ensemble model of decision tree (DT) and reduced error pruning (REP) algorithms, which is equally effective for classification and regression problems [19]. This algorithm uses knowledge gain to construct a decision tree and prunes it using reduced-error pruning. Since complex decision trees can lead to overfitting and reduced model interpretability, REP reduces complexity by eliminate leaves and branches from the DT structure.

3.4.5. J48

The J48 has features such as missing values, rule derivation, continuous attribute value ranges, and decision tree pruning, among others. If possible, overfitting pruning may be used as a precision device [20]. This algorithm used to the creation of the rules for lead to the formation of a unique identity for the data. The aim of using this classifier is to gradually mainstream the decision tree until it achieves a balance between versatility and accuracy.

3.4.6. Random Forest (RF)

Random Forest Classification (RFC) is a supervised classification technique for machine learning that is based on decision trees [21].

Each tree in the collection is created by choosing a small group of features to split on at random for each node, and then determining the best split based on these features in the training set. Each tree assigns a vote to that particular feature vector. For each feature vector, the forest chooses the class with the most votes. It is simple to create and forecast, runs quickly on large datasets, easily estimates missing data, and retains accuracy even when a large percentage of data is missing.

3.4.7. k-nearest neighbor

The KNN is a system that is based on supervised learning, which allows machines to categorize objects, problems, or circumstances based on data that has already been fed into them [22]. k is a user-defined constant in the classification process, and an unlabeled vector (a query or test point) is categorized by assigning the mark that appears most frequently among the k training samples closest to that query point. Euclidean distance was used as a distance metric for continuous variables, this metric calculated according to Eq. (5) below:

$$d(x,y) = \sqrt{\sum_{i=1}^{n}(a_i(x) - a_i(y))^2} \quad [22]$$

4. The proposed system implementation

This system was implemented into two Phases as follows:

4.1. Training phase

The proposed system's first step is the training of the two datasets using cross-validation, where largest part of data enters to training phase then the reminder data passed to testing phase sec. (4.2). The dataset will be preprocessed using a hamming window, after which the features will be extracted and modeled using MFCC and VQ, and the values of these features will be combined with the features extracted to prepare for the classification algorithms. The mixed features are saved as reference models during the training process. These models are then compared to the speech signals that have been entered.

4.2. Testing phase

The proposed system's testing phase is the second phase. As mentioned above the reminder data will be tested after applying the same pre-processing steps which applied on data in training phase. The proposed system architecture is as shown in Algorithm (1) below.

Algorithm (1) The proposed voice recognition system

Input -: Voice datasets

Output: - Best classifier performance

Begin

// **Input** //

1: Load dataset (1), (2)

2: Removing noise using Hamming window

3: Make the spectral form of the speech signal frequency more smooth using Pre-emphasis

Pre-processing Phase

4: Signals converting into a frequency spectrum using Fast Fourier Transform (FFT)

5: Human hearing activity to detect frequencies using Mel-Frequency Cepstral Coefficients (MFCC)

Feature Extraction Phase

6: Digital representation of signals using Vector Quantization (VQ)

7: Ruffle the dataset randomly and divide it into (k) groups using Cross-Validation, a (k –1) sub-instances were applied for training

8: The remaining data single sub-instance will be applied as the validation data for testing

9: Classify instances based 7 Classifiers

10: Classifiers Evaluation

11: Best Classifier performance (Accuracy, Precision, Recall, F- measure, Error Rate, Kappa) **Output**

End

Table 2
Results of Machine Learning Classifiers.

	PART		JRip		NB		RT		J48		RF		KNN	
	Dataset1	Dataset2	Dataset1	Dataset2	Dataset1	Dataset2	Dataset1	Dataset2	Dataset1	Dataset2	Dataset1	Dataset2	Dataset1	Dataset2
Total instances	7500	2226	7500	2226	7500	2226	7500	2226	7500	2226	7500	2226	7500	2226
Total correct	3879	529	7033	1740	6818	2093	6920	1858	6961	1940	7352	2140	6766	415
Total incorrect	362	1697	467	486	682	133	580	368	539	286	148	86	734	1811
Accuracy	0.517	0.237	0.937	0.781	0.909	0.94	0.929	0.834	0.928	0.871	0.988	0.961	0.902	0.186
Precision	0.51	0.18	0.93	0.81	0.91	0.95	0.92	0.84	0.92	0.87	0.98	0.96	0.91	0.07
Recall	0.51	0.23	0.93	0.78	0.90	0.94	0.92	0.83	0.92	0.87	0.98	0.96	0.90	0.18
F- measure	0.51	0.19	0.93	0.78	0.91	0.94	0.92	0.83	0.92	0.87	0.98	0.96	0.90	0.09
Error rate	0.482	0.762	0.622	0.218	0.091	0.06	0.071	0.166	0.072	0.129	0.012	0.039	0.098	0.814
Specificity	0.87	0.97	0.98	0.99	0.97	0.99	0.98	0.99	0.98	0.99	0.99	0.99	0.97	0.96
KAPPA	0.39	0.211	0.92	0.77	0.886	0.93	0.903	0.82	0.91	0.86	0.97	0.96	0.87	0.15

5. Proposed system evaluation

For evaluating a model's performance, certain parameters are used to determine its behavior. The results are influenced by the size of the training data, the quality of the audio files, and, most importantly, the type of machine-learning algorithm used. The following criteria are used to assess the models' efficacy [23]:

- **Accuracy:** Percentage of examples correctly categorized from all given examples. It is calculated as:

$$\mathrm{Accuracy} = \frac{\mathrm{tp} + \mathrm{tn}}{\mathrm{tp} + \mathrm{tn} + \mathrm{fp} + \mathrm{fn}}[23] \tag{6}$$

- **Precision:** The percentage of true x-class instances for all those listed as class x. It is calculated as:

$$\mathrm{Precision} = \frac{\mathrm{tp}}{\mathrm{tp} + \mathrm{fp}}[23] \tag{7}$$

- **Recall:** The percentage of examples listed as class $\times$ among all examples of class x. It is calculated as:

$$\mathrm{Recall} = \frac{\mathrm{tp}}{\mathrm{tp} + \mathrm{fn}}[23] \tag{8}$$

- **F- measure:** is the harmonic mean of precision and recall. It is calculated as:

$$F_1 = 2 * \frac{\mathrm{precision} * \mathrm{recall}}{\mathrm{precision} + \mathrm{recall}}[23] \tag{9}$$

where

tp = true positives: number of examples predicted positive that are actually positive
fp = false positives: number of examples predicted positive that are actually negative
tn = true negatives: number of examples predicted negative that are actually negative
fn = false negatives: number of examples predicted negative that are actually positive

- **Error rate:** An error is simply a misclassification: the classifier has presented a case, and it classifies the case incorrectly, as shown in Eq. (10) below:

$$\mathrm{ErrorRate} = 1 - \mathrm{accuracy}[23] \tag{10}$$

- **Specificity:** measures the ability of a test to be negative when the condition is actually not present. It is also known as false-positive rate, precision, Type I error, α error, the error of commission, or null hypothesis.

$$\mathrm{Specificity} = \frac{\mathrm{TN}}{\mathrm{TN} + \mathrm{FP}}100\%[23] \tag{11}$$

- **KAPPA:** Cohen's kappa coefficient can be applied for evaluating agreement between two regular nominal classifications. If one uses Cohen's kappa to quantify balance between the classifications, the ranges between all categories are considered identical, and this makes sense if all nominal categories reflect different kinds of 'presence' [24]. The weighted kappa coefficient is defined as:

$$K = \frac{O - E}{I - E}[24] \tag{12}$$

6. Experiential results

In this experiment, we test our datasets with (PART, JRip, NB, RT, J48, RF, KNN) classifiers. The results of using the voice (dataset1, dataset2) as input will illustrate in Table 2. The first dataset contains five persons and 1500 samples for each one of them. From Table 2 and figures numbered from Figs. 3–9, for dataset1 can notice that the Random Forest (RF) classifier gives the best accuracy, precision, recall, f-measure and specificity while PART classifier gives the worst accuracy, precision recall, f-measure and specificity. As for error rate the results were on the contrary, where the PART classifier produced highest error rate and KNN produced

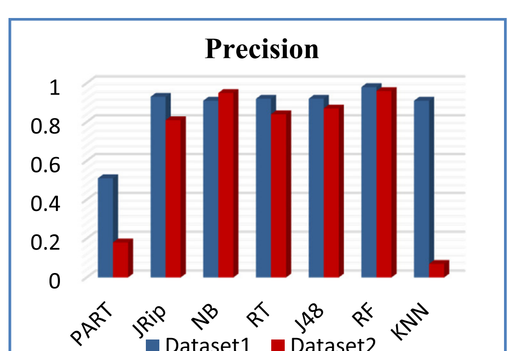

Fig. 3. Precision measured for Classifiers.

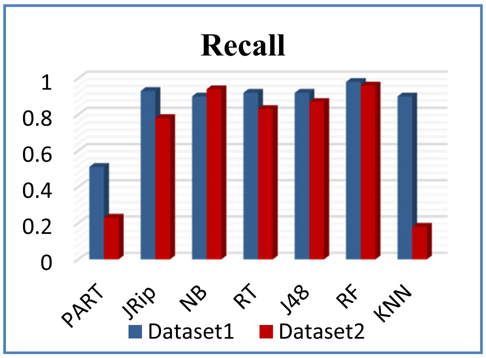

Fig. 4. Recall measured for Classifiers.

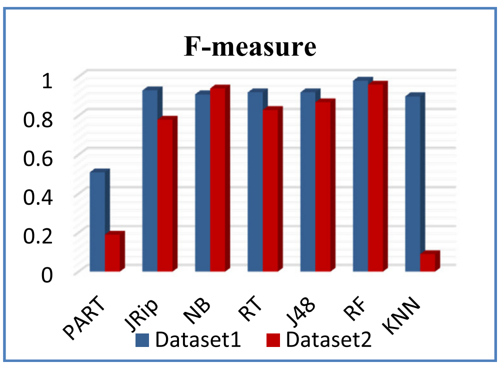

Fig. 5. F-measure measured for Classifiers.

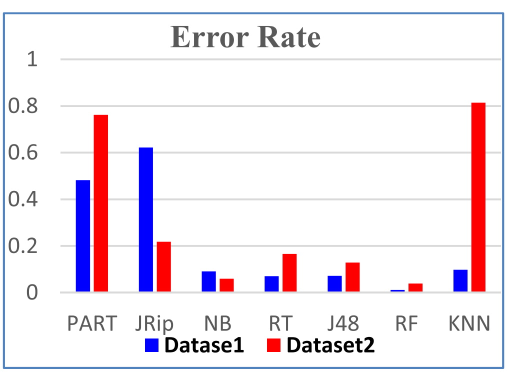

Fig. 6. Error Rate measured for Classifiers.

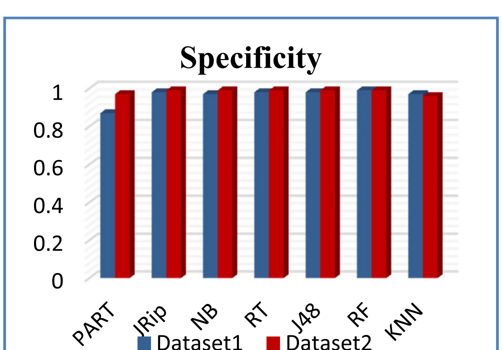

Fig. 7. Specificity measured for Classifiers.

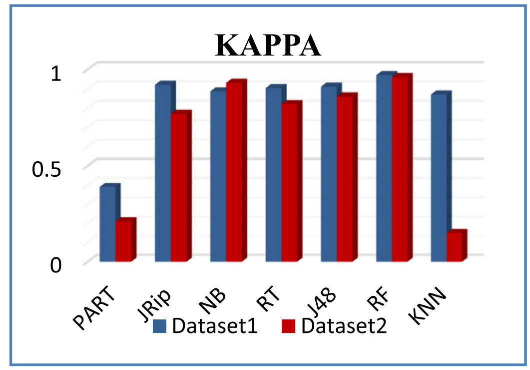

Fig. 8. KAPPA measured for Classifiers.

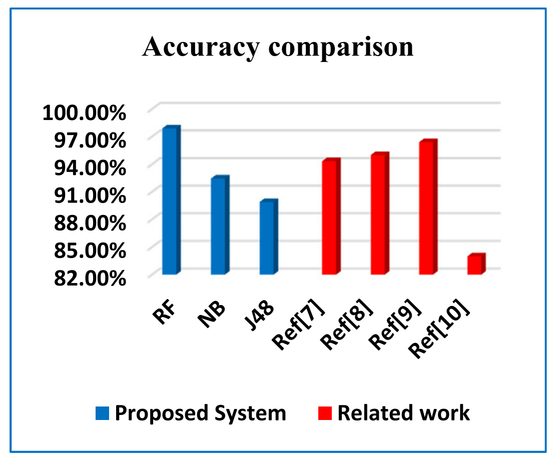

Fig. 9. Results Comparison Figure.

Table 3
Results Comparison.

	RF	NB	J48	Ref. [7]	Ref. [8]	Ref. [9]	Ref. [10]
Accuracy	97.9%	92.45%	89.9%	94.33%	95%	96.4%	84%

lowest error rate among other classifiers. The second dataset contains 50 persons and 1500 samples for each one of them. With this dataset can notice that the RF Classifier gives the best accuracy, precision, recall and f-measure while KNN classifier gives the worst accuracy precision, recall and f-measure. As for error rate the results were on the contrary, where the KNN classifier produced highest error rate and RF produced lowest error rate among other classifiers (Table 3).

7. Results comparison

In Table 3 the comparison between the best three performances of proposed system algorithms with the related work. The average accuracy of the proposed system for the three classifiers (RF,J48, NB) are calculated by Eq. (13):

$$\text{ave.accuracy} = \frac{accuracy\ of\ dataset1 + accuracy\ of\ dataset2}{2} \quad (13)$$

The results refer to that the proposed system showed superiority using Random Forest algorithm in terms of the average accuracy of voice recognition over all Literature Survey techniques which mentioned in Table 3

8. Conclusions

Audio preprocessing, feature extraction, and machine learning classification are the three main components of this study. Since the audios used in our study were not captured in confined spaces, audio pre-processing was a critical component of the study. Reduced ambient noise and emphasizing human vocals were the two most critical aspects we focused on for pre-processing. We thought that relying solely on the MFCC coefficients would be sufficient for the analysis. Two datasets were used and seven machine learning algorithms. Our findings showed that using a machine learning classifier in the classification process increased accuracy, with (RF) classifiers reaching 97.9% accuracy. It was superior to the accuracy results of the previous work, noting that the tested data sets differed

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] M. Algabri, H. Mathkour, M.A. Bencherif, M. Alsulaiman, M.A. Mekhtiche, Automatic speaker recognition for mobile forensic applications, Mobile Information Systems 2017 (2017).
[2] Anggraeni, D., W. S. M. Sanjaya, M. Y. S. Nurasyidiek, and M. Munawwaroh. ''the implementation of speech recognition using mel-frequency cepstrum coefficients (MFCC) and support vector machine (SVM) method based on python to control robot arm.'' In IOP Conference Series: Materials Science and Engineering, vol. 288, no. 1, p. 012042. IOP Publishing, 2018.
[3] Hasan, Md Al Mehedi, and Shamim Ahmad. ''PredSucc-site: Lysine succinylation sites prediction in proteins by using support vector machine and resolving data imbalance issue.'' International Journal of Computer Applications 182, no. 15 (2018): 8887.
[4] ELLaban, Hend Ab, Ahmed A. Ewees, and Abdelrazek E. Elsaeed. ''A real-time system for facial expression recognition using support vector machines and k-nearest neighbor classifier.'' International Journal of Computer Applications 159, no. 8 (2017): 23-29.
[5] Saleh Khawatreh, Belal Ayyoub, Ashraf Abu-Ein, Ziad Alqadi, A Novel Methodology to Extract Voice Signal Features, International Journal of Computer Applications 975 (2018) 8887.
[6] Tayseer MF Taha, Ahsan Adeel, Amir Hussain, A survey on techniques for enhancing speech, International Journal of Computer Applications 179 (17) (2018) 1–14.
[7] Suma Paulose, Dominic Mathew, Abraham Thomas, Performance evaluation of different modeling methods and classifiers with MFCC and IHC features for speaker recognition, Procedia Comput. Sci. 115 (2017) 55–62.
[8] R. Thiruvengatanadhan, Speech Recognition using SVM, International Research Journal of Engineering and Technology (IRJET) 5, no. 09 (2018).
[9] Nassren A. Alwahed, Talib M. Jawad, ARABIC SPEECH RECOGNITION BASED ON KNN, J48, AND LVQ, Iraqi Journal of Information & Communications Technology 2 (2) (2019) 1–8.
[10] M. Subba, G. Lakshmi, P. Gowri and K. Chowdary ''RANDOM FOREST BASED AUTOMATIC SPEAKER RECOGNITION SYSTEM.'' The International Journal of analytical and experimental modal analysis, pp: 526- 535, April/2020.
[11] Nidaa F. Hassan, Sarah Qusay Selah Alden, Gender classification based on audio features, Journal of Al-Ma'moon College 31 (2018).
[12] Sullivan, Michael. ''Global markets and technologies for voice recognition.'' Information Technology Market Research Reports in BCC Research (2017).
[13] Seyedmahdad Mirsamadi, Emad Barsoum, Ch.a. Zhang, in: Automatic speech emotion recognition using recurrent neural networks with local attention, IEEE, 2017, pp. 2227–2231.
[14] Z. Aldeneh and E. Provost, ''Using regional saliency for speech emotion recognition.'' in Acoustics,Speech, and Signal Processing (ICASSP), 2017 IEEE International Conference on. IEEE, 2017, pp. 2741–2745.
[15] Ranjodh Singh, Hemant Yadav, Mohit Sharma, Sandeep Gosain, Rajiv Ratn Shah, in: Automatic Speech Recognition for Real Time Systems, IEEE, 2019, pp. 189–198.
[16] Amandeep Kaur Sandhu, Ranbir Singh Batth, Software reuse analytics using integrated random forest and gradient boosting machine learning algorithm, Software: Practice and Experience 51 (4) (2021) 735–747.
[17] Parashar Dhakal, Praveen Damacharla, Ahmad Y. Javaid, Vijay Devabhaktuni, in: Detection and Identification of Background Sounds to Improvise Voice Interface in Critical Environments, IEEE, 2018, pp. 078–083.
[18] Parashar Dhakal, Praveen Damacharla, Ahmad Y. Javaid, Vijay Devabhaktuni, A near real-time automatic speaker recognition architecture for voice-based user interface, Machine Learning and Knowledge Extraction 1 (1) (2019) 504–520.
[19] Samira Hazmoune, Fateh Bougamouza, Smaine Mazouzi, Mohamed Benmohammed, A new hybrid framework based on Hidden Markov models and K-nearest neighbors for speech recognition, Int. J. Speech Technol. 21 (3) (2018) 689–704.
[20] Peng Song, Wenming Zheng, Feature selection based transfer subspace learning for speech emotion recognition, IEEE Trans. Affective Comput. 11 (3) (2018) 373–382.
[21] Joeky T. Senders, Mark M. Zaki, Aditya V. Karhade, Bliss Chang, William B. Gormley, Marike L. Broekman, Timothy R. Smith, Omar Arnaout, An introduction and overview of machine learning in neurosurgical care, Acta Neurochir. 160 (1) (2018) 29–38.
[22] Mohammad Gohari, Amir Mohammad Eydi, Modelling of shaft unbalance: Modelling a multi discs rotor using K-Nearest Neighbor and Decision Tree Algorithms, Measurement 151 (2020) 107253.
[23] Ali Hamid Meftah, Yousef Ajami Alotaibi, Sid-Ahmed Selouani, ''Evaluation of an Arabic speech corpus of emotions: A perceptual and statistical analysis.'' IEEE, Access 6 (2018) 72845–72861.
[24] Matthijs J. Warrens, Kappa coefficients for dichotomous-nominal classifications, Adv. Data Anal. Classif. (2020) 1–16.

Further reading

[25] B. ALhayani, and H. Ilhan, ''Efficient cooperative imge transmission in one-Way mult-hop sensor network,'' International Journal of Electrical Engineering Education, vol.57, no.2, 321-339. 2020.
[26] E. Milind, Rane, Umesh S Bhadade, ''Comparative Study of ROI Extraction of Palmprint'', IJCSN International Journal of Computer Science and Network 5 (2) (April 2016).
[27] Milind. Rane and Umesh. Bhadade, '' Multimodal score level fusion for recognition using face and palmprint'', The International Journal of Electrical Engineering & Education, PP1-19, 2020
[28] Milind Rane, Tejas Latne, Umesh Bhadade, Biometric Recognition Using Fusion, ICDSMLA 1320–1329 (2019) 2019.
[29] Alhayani, B.S.A., llhan, H. Visual sensor intelligent module based image transmission in industrial manufacturing for monitoring and manipulation problems. J Intell Manuf 32, 597–610 (2021). 10.1007/s10845-020-01590-1
[30] B. Alhayani, A.A. Abdallah, Manufacturing intelligent Corvus corone module for a secured two way image transmission under WSN, Engineering Computations 37 (9) (2020) 1–17.
[31] B. ALhayani and H. Ilhan, ''Image transmission over decode and forward based cooperative wireless multimedia sensor networks for Rayleigh fading channels in medical internet of things (MIoT) for remote health-care and health communication monitoring,'' Journal of Medical Imaging And Health Informatics, vol. 10, pp. 160-168.2020
[32] B. .Alhayani and Milind Rane,''face recognition system by image processing'' International journal of electronics and communication engineering & technology (IJCIET),vol.5, no.5, 80–90. 2014.
[33] Bilal Alhayani, Husam Jasim Mohammed, Ibrahim Zeghaiton Chaloob, Jehan Saleh Ahmed, Effectiveness of artificial intelligence techniques against cyber security risks apply of IT industry, Mater. Today:. Proc. (2021), https://doi.org/10.1016/j.matpr.2021.02.531.
[34] Bilal Alhayani, Sara Taher Abbas, Dawood Zahi Khutar, Husam Jasim Mohammed, Best ways computation intelligent of face cyber attacks, MaterialsToday Proceedings (2021), https://doi.org/10.1016/j.matpr.2021.02.557.