

Voice Disorder Identification by using Machine Learning Techniques

Laura Verde, Giuseppe De Pietro, *Member, IEEE*, and Giovanna Sannino, *Member, IEEE*

Abstract—Nowadays, the use of mobile devices in the health-care sector is increasing significantly. Mobile technologies offer not only forms of communication for multimedia content (e.g. clinical audio-visual notes and medical records) but also promising solutions for people who desire the detection, monitoring and treatment of their health conditions anywhere and at any time. Mobile health systems can contribute to make patient care faster, better and cheaper. Several pathological conditions can benefit from the use of mobile technologies. In this work we focus on dysphonia, an alteration of the voice quality that affects about one person in three at least once in his/her lifetime. Voice disorders are rapidly spreading, although they are often underestimated. Mobile health systems can be an easy and fast support to voice pathology detection. The identification of an algorithm that discriminates between pathological and healthy voices with more accuracy is necessary to realise a valid and precise mobile health system. The key contribution of this study is to investigate and compare the performance of several machine learning techniques useful for voice pathology detection. All analyses are performed on a dataset of voices selected from the Saarbruecken Voice Database. The results obtained are evaluated in terms of accuracy, sensitivity, specificity and ROC area. They show that the best accuracy in voice diseases detection is achieved by the Support Vector Machine algorithm or the Decision Tree one, depending on the features evaluated by using opportune feature selection methods.

Index Terms—Mobile Health Systems, Machine Learning Techniques, Voice Disorders, Classification Accuracy.

I. INTRODUCTION

The introduction of mobile devices for data transmission or disease control and monitoring has been a main attraction of research and business communities. They offer, in fact, numerous opportunities to realise efficient mobile health (m-health) systems. These solutions can allow patients and doctors to access medical records, clinical audio-visual notes and drug information anywhere and at any time from their mobile devices, such as a tablet or smartphone, to monitor several conditions [1]. M-health solutions can also be used in other important applications such as the detection and prevention of specific diseases, decision making and the management of chronic conditions and emergencies, improving the quality of patient care and reducing the costs of healthcare.

Several pathological conditions can be detected and monitored, such as the well known and widespread cardiovascular

diseases. In recent years, probably also due to the diffusion of the Internet of Things (IoT) and cloud technologies, there has been a development of monitoring systems in an unobtrusive, portable and easy way using wearable sensors and wireless communications, such as the solutions described in [2], [3], [4], [5], [6], [7]. These systems are able to achieve health data monitoring and analysis, helpful for patients suffering from cardiovascular diseases or for their physical therapy.

If, on the one hand, health monitoring systems for cardiovascular diseases are so celebrated, on the other hand, there are other little known and often underestimated disorders, such as dysphonia, that could benefit from m-health solutions.

Dysphonia is a disorder that occurs when the voice quality, pitch and loudness are altered. About 10% of the population suffer from this disorder [8], caused mainly by unhealthy social habits and voice abuse. Unfortunately, a large number of individuals with voice disorders do not seek treatment. Therefore, m-health systems could be an efficient support for the diagnosis and screening of voice disorders.

Clinical voice pathology detection is performed through the execution of several procedures, such as the acoustic analysis. It consists of an estimation of appropriate parameters extracted from voice signal to evaluate any possible alterations of the vocal tract, according to the guidelines of the SIFEL protocol [9] (Società Italiana di Foniatria e Logopedia), developed by the Italian Society of Logopedics and Phoniatrics, following the instructions of the Committee for Phoniatrics of the European Society of Laryngology. It is a non-invasive examination in clinical practice, complementary to other medical tests, such as the laryngoscopic examination based on the direct observation of the vocal folds.

Several acoustic parameters are estimated to evaluate the state of health of the voice. Unfortunately, the accuracy of these parameters in the detection of voice disorders is, often, related to the algorithms used to estimate them. For this reason the main effort of researchers is oriented to the study of acoustic parameters and the application of classification techniques able to obtain a high discrimination accuracy. Recently, speech pathology has focused interest on machine learning techniques.

In this work, we want to discuss the application of machine learning algorithms and features selection methods capable of discriminating between pathological and healthy voices with a better accuracy. In detail, we evaluate the pathology recognition using the information data of patients, such as age and gender, and different features extracted from the voice signals. The adopted parameters are those estimated in the clinical acoustic analysis, such as the Fundamental Frequency

Laura Verde is with the Department of Engineering, University of Naples Parthenope, Centro Direzionale di Napoli, Isola C4 - Naples, Italy; e-mail: laura.verde@uniparthenope.it.

Giuseppe De Pietro and Giovanna Sannino are with the Institute of High Performance Computing and Networking (ICAR-CNR), Via Pietro Castellino, 111 - Naples, Italy; e-mail: giuseppe.depietro@icar.cnr.it (Giuseppe De Pietro) and giovanna.sannino@icar.cnr.it (Giovanna Sannino).

TABLE I
SUMMARY OF RESEARCH STUDIES ON VOICE PATHOLOGY DETECTION.

Ref.	Database used	#of voices (pathological + healthy)	#of voices used for training (pathological + healthy)	#of voices used for testing (pathological + healthy)	Features used	Classifier used	Accuracy(%) over the testing set	Accuracy(%) over the training set
[10]	MEEI	226 (173+53)	158 (n.a.)	68 (n.a.)	MFCC, Noise Features and Temporal Derivatives	SVM	95	n.a.
[11]	SVD	110 (40+70)	77 (n.a.)	33 (n.a.)	MFCC and Temporal Derivatives	SVM	86	n.a.
[12]	MEEI	154 (118+36)	4-fold	classification	MFCC	GMM	77.90	n.a.
[13]	SVD	100 (40+60)	75 (35+40)	25 (10+15)	MFCC	SVM, GMM	96.5 - 95.5	n.a.
[14]	SVD	101 (38+63)	81 (29+52)	20 (9+11)	MFCC, jitter and shimmer	GMM	82.37	n.a.
[15]	Private	60 (30+30)	48 (n.a.)	12 (n.a.)	Wavelet Transform	LS-SVM	90	n.a.
[16]	Private	152 (111+41)	101 (n.a.)	51 (n.a.)	jitter, shimmer, NHR, SPI, APQ and RAP	GMM	95.2	98.4
[17]	Private	77 (n.a.)	n.a.	n.a.	Spectral Features	GMM	94	n.a.
[18]	Private	400 (300+100)	200 (150+50)	200 (150+50)	Perturbation, noise and energy parameters	K-NN, LDA	93.5	n.a.

(F_0), jitter, shimmer and Harmonic to Noise Ratio (HNR). In addition, other parameters, the Mel-Frequency Cepstral Coefficients (MFCC), the first and second derivatives, are used due to their wide application both in machine learning techniques and in voice disorders classification as reported in several studies [19], [20]. The performances are evaluated in terms of accuracy, sensitivity, specificity and ROC area for each considered machine learning methods.

The remaining sections of the paper are organized as follows. In Section II, we present the main studies existing in literature concerning the use of machine learning techniques to estimate voice disorders. In Section III, we introduce the experimental phase performed in this study, focusing on the dataset, features and machine learning algorithms used for the classification. The results obtained are discussed in the Section IV, while our conclusions are provided in Section V.

II. RELATED WORK

Speech or, in general, the voice signal is used in several kinds of application ranging from emotion recognition [21] to patient healthcare state recognition [22].

Several m-health solutions, such as [23], [24], [25], adopt these signals to estimate the state of voice health, as well as systems that use voice signals to evaluate emotional condition [26], [27]. Voice pathology detection has, often, been achieved through specific machine learning techniques, and over recent years, several approaches have been developed to improve the performance in terms of accuracy in the discrimination between healthy and pathological voices.

These studies are focused on the identification of parameters to measure the voice quality and new techniques able to detect voice disorders.

Among several machine learning techniques existing in literature, Support Vector Machine (SVM) has been widely used in voice signal processing. Godino et al. [10], for example,

focused on the classification of pathological and healthy voices based on MFCC to train and test an SVM classifier. These have obtained a good accuracy (95%). However, the poor numerosity of the used dataset composed of only 173 pathological and 53 healthy voices selected by the Massachusetts Eye and Ear Infirmary voice and speech lab (MEEI) database [28] should be observed. Additionally, important information, as for example the pathologies of the selected voices, is not available in this work.

The SVM technique was also used in [11] to estimate the presence of dysphonia, investigating four types of pathology: chronic laryngitis, cysts, Reinke's edema and spasmodic dysphonia. The authors proposed an algorithm based on the use of MFCC and Linear Discriminant Analysis (LDA) as a dimensionality reduction method. This algorithm identifies the presence of a pathology with a discrete accuracy (86%). However, it was tested on a very limited dataset. In fact, only 70 pathological and 40 healthy voices were selected by the Saarbruechen Voice Database (SVD) [29].

The MFCC parameters were considered in other numerous studies, such as [12], [13]. In [12] subjects with nodules, edema and unilateral vocal fold paralysis were analysed with not so encouraging results (77.90%), while in [13] patients suffering from spasmodic dysphonia were selected. Unfortunately, the performance of the algorithms, Gaussian Mixture Models (GMM) and SVM, was tested on a limited datasets composed of voices extracted from the MEEI database.

El Emary et al. [14], instead, classified the speech signal by estimating not only the MFCC but also jitter and shimmer. The detection of voice suffering from neurological disorders was performed using the GMM algorithm on a very small subset of the SVD database containing only 38 pathological and 63 healthy voices. In [15], an algorithm based on Daubechies discrete wavelet transform, linear prediction coefficients and last squares support vector machine (LS-SVM) was used to



Fig. 1. The flowchart of a possible m-Health system for the voice health state classification.

identify laryngeal pathologies. The experiments were carried out using a private database.

Another private dataset collected in the Busan National University Hospital was used in the study described in [16]. Wang et al. classified pathological voices using Hidden Markov Models (HMM), GMM and SVM. The voice disorders considered in this study included vocal polyps, vocal cord palsy, nodules, cysts, edema, laryngitis and glottic cancer. In several studies existing in literature private databases were used, such as for example in [17], [18]. In the first case the data adopted to test the developed system were captured at the Christie and Withington Hospitals in Manchester. Only 77 abnormal speech signals were used to train and test the proposed artificial neural network, and the authors did not specify the pathologies considered. In [18] the voices of the subjects were collected at the Phoniatic Department of the University Hospital of Sofia to detect people suffering from laryngeal pathology via the K-nearest neighbours algorithm and linear discriminant analysis.

Table I summarizes the different approaches outlined so far. Most of the features and algorithms were trained and tested using limited databases, including a few types of disorders and a few voices. In many cases, the databases used are not available and their results cannot be compared. Moreover, the authors often use only the MFCC as the signal features, not considering the characteristic parameters indicated by clinical protocols, such as the SIFEL protocol, to evaluate the voice quality and the possible presence of disorders.

III. MATERIALS AND METHODS

In this study we analysed the accuracy in the discrimination of pathological from healthy voices of the main machine learning techniques to identify the most reliable one. The idea has been to integrate the best one in a valid m-health system, where the voice signal can be acquired by a mobile device, such as a smartphone or tablet, processed in real-time to extract the voice features, and analyzed by using the machine learning classifier to detect the presence or not of a voice disorders, as shown in Figure 1.

In detail, we have evaluated the performance of SVM, the principal adopted technique in literature in relation to the Kernel function, and of some other machine learning algorithms used to identify the presence of voice disorders.

The analysis has been performed using the WEKA [30] tool, one of the most commonly used tools for data mining tasks, selected for the data analysis due to its efficiency, versatility and affordability.

In the following subsections we introduce the dataset used in this study, the features extracted from the voice signal and used for the classification, and the machine learning techniques compared.

A. The database

In our research, we have selected a subset of voice samples from the "Saarbruecken Voice Database" (SVD) [31], [32]. SVD database is a collection of 2041 voice recordings, containing voices from healthy and pathological individuals, published online by the Institute of Phonetics of the University of Saarland. All the recordings are sampled at 50 kHz and their resolution is 16-bit. In total, there are 1354 pathological voices (627 male and 727 female), suffering from 71 different diseases, distinguished between functional and organic disorders. The remaining 687 healthy voices are 259 male and 428 female.

This collection consists of recordings of vowels /a/, /i/, /u/ and an appropriate sentence. To evaluate the patient's voice quality the use of vowels is preferable because they avoid linguistic artifacts and are used in many voice assessment applications [33]. In relation to voice disorder detection and identification problems, in clinical practice, the vocalization of the vowel /a/ is used.

In our experimental tests, to perform the experiments on a well-balanced database containing both pathological and healthy voices, we have selected a total of 1370 /a/ vocalizations. In detail, we have chosen:

- 685 pathological voices (257 male and 428 female); and
- 685 healthy voices (257 male and 428 female).

More details of the selected recordings are indicated in Table II, in which we have reported the number (No) of considered voices for each age and gender, and the percentage calculated in relation to the whole dataset. The lower number of male samples than female ones is related to the higher incidence of voice disorders in female subjects than in males [34]. We have used all the available healthy and pathological voices from the SVD.

All samples contain the recording of the vowel /a/, the signal required by the SIFEL protocol to evaluate the voice's state of health [9]. To test the capability of the considered algorithms, we have selected the pathological voices from all types of pathology existing in the database. There are organic voice disorders, such as chronic laryngitis or Reinke's edema, and functional dysphonia as a hyperfunctional or hypofunctional one.

B. Features used for the classification

Feature extraction is an important task that allows an improvement of the analysis and classification. The choice of which features of the speech signal to use in our study was made by taking into account two considerations. On the one hand, we have used the main parameters adopted by the specialist during the clinical evaluation; on the other, we have

TABLE II
DETAILS OF THE VOICE SIGNALS USED IN THIS STUDY.

Category	Gender	Age Group	No.	%
Healthy	Female	17-29	359	26.20 %
		30-39	27	1.98 %
		40-49	15	1.09 %
		50-59	13	0.95 %
		60+	14	1.02 %
Healthy	Male	17-29	138	10.07 %
		30-39	62	4.53 %
		40-49	37	2.70 %
		50-59	9	0.66 %
		60+	11	0.80 %
Pathological	Female	17-29	58	4.24 %
		30-39	94	6.86 %
		40-49	85	6.20 %
		50-59	87	6.35 %
		60+	104	7.59 %
Pathological	Male	17-29	23	1.68 %
		30-39	24	1.75 %
		40-49	43	3.14 %
		50-59	74	5.40 %
		60+	93	6.79 %
Total	<i>Female</i>	<i>17-60+</i>	<i>856</i>	<i>62.48%</i>
	<i>Male</i>	<i>17-60+</i>	<i>514</i>	<i>37.52%</i>

chosen the main features used in several correlated studies existing in literature concerning the use of machine learning techniques for the voice classification.

In detail, the parameters used in clinical practice are:

- **Fundamental Frequency (F_0):** this represents the rate of vibration of the vocal folds constituting an important index of laryngeal function. It is at the basis of the other parameters calculated in the acoustic analysis and most noise estimation methods [18], [35].
- **Jitter:** this describes the instabilities of the oscillating pattern of the vocal folds, quantifying the cycle-to-cycle changes in fundamental frequency.
- **Shimmer:** this indicates the instabilities of the oscillating pattern of the vocal folds, quantifying the cycle-to-cycle changes in amplitude.
- **Harmonic to Noise Ratio (HNR):** this quantifies the ratio of signal information over noise due to turbulent airflow, resulting from an incomplete vocal fold closure in speech pathologies.

The parameters used in other correlated studies are:

- **Mel-Frequency Cepstral Coefficients (MFCC):** these coefficients try to analyse the vocal tract independently of the vocal folds that can be damaged due to voice pathologies. In this work, the experiments were conducted using 13 MFCC coefficients.
- **First and second derivatives of cepstral coefficient:** these are useful to investigate the properties of the dynamic behaviour of the speech signal.

It is important to note that, for some of the above mentioned features, there are no standard algorithms available for their calculation. This is a critical issue, because the more accurate is the computation of each parameter, the more

reliable is the voice analysis, namely the classification of the voice signal as healthy or pathological. For example, for the evaluation of the F_0 , there are several different methods proposed in literature, like Spectral Analysis [36], the Hilbert-Huang transform[37], the Robust Algorithm for Pitch Tracking (RAPT) [38], the Dynamic Programming Projected Phase-Slope Algorithm (DYPSA) [39], the Speech Transformation and Representation based on Adaptive Interpolation of weiGHTed spectrogram (STRAIGHT) method [40] and the extraction based on the Autocorrelation Function of the speech signal [41].

However, in our study, we have used a new proposed methodology, documented in [42], that is an optimization and personalization of the Yin algorithm [43].

Concerning the jitter and shimmer features, they were estimated utilizing the method presented in [44]. In detail, the jitter was expressed as a percentage and it was calculated as the average absolute difference between consecutive periods, divided by the average period, that is:

$$Jitter(\%) = \frac{\frac{1}{N-1} \sum_{i=1}^{N-1} (|T_i - T_{i+1}|)}{\frac{1}{N} \sum_{i=1}^N (T_i)} \quad (1)$$

where T_i are the consecutive periods and N is the number of extracted F_0 periods.

The shimmer, instead, was estimated as the average absolute base-10 logarithm of the difference between the amplitudes of consecutive periods, multiplied by 20. It was expressed in decibels (dB) by the following expression:

$$Shimmer(dB) = \frac{1}{N-1} \sum_{i=1}^{N-1} |20 \log \frac{A_{i+1}}{A_i}| \quad (2)$$

where A_i are the extracted peak-to-peak amplitudes and N is the number of extracted F_0 periods.

The HNR was computed in dB as the mean difference between the harmonic peaks and the aperiodic components according to de Krom's algorithm [45].

Finally, the MFCC coefficients resulting from by the cepstral representation of the voice signal, were calculated by evaluating the Discrete Cosine Transform and the log compression of the voice samples in the frequency domain. These coefficients and their derivatives were extracted using the *melcepst* Matlab function of the VOICEBOX [46] tool, a speech processing toolbox realised by the Department of Electrical and Electronic Engineering of Imperial College of London and used by several studies existing in literature such as [13] and [47].

In summary, each instance i of the database used in this study is constituted by the following information:

- **Subject ID:** a number value to identify the subject;
- **age:** measured by years from birth;
- **gender:** female or male;
- **features:** F_0 , jitter, shimmer, HNR, MFCC (from 1 to 13), first derivative and second derivative, calculated over the recording of the vowel /a/;
- **class:** pathological or healthy.

C. Machine Learning Classifiers

In order to make an exhaustive comparison, we have chosen different machine learning algorithms. Actually, each of them has been chosen as a representative of a class of algorithms based on similar characteristics. These techniques are:

- **Support Vector Machine (SVM):** this is a discriminative classifier formally defined by a separating hyperplane that divides data belonging to different classes. The aim is to identify the class of belonging of the different data. Training a support vector machine requires the solution of a very large quadratic programming optimization problem. To resolve this problem the sequential minimum optimization (SMO) technique is used, which is able to divide the optimization problem into a series of smaller possible problems [48]. The classification accuracy can be improve by selecting opportune form and parameters characteristic of the kernel function $K(x,y)$ [10], [49]. The most popular kernel function forms are polynomial and radial basis ones [50].
- **Decision Tree (DT):** this technique is used to classify categorical data in which the learned function is represented by a decision tree. Decision trees are easy to interpret, capable of working with missing values and categorical and continuous data, characteristics of the medical field. We have used J48, an implementation of algorithm C4.5 [51], the most popular tree classifier.
- **Bayesian Classification (BC):** this approach named after Thomas Bayes, who proposed the Bayes Theorem. The classification is achieved by evaluating the probabilistic model that represents a set of random variables and their conditional dependencies identified, respectively, as nodes and strings [52]. The major advantage is the easy interpretation of the results and the robustness in dealing with missing data.
- **Logistic Model Tree (LMT):** this technique combines logistic regression models with tree induction. It consists of a standard decision tree structure with logistic regression functions at the leaves. SimpleLogistic class [53] implements this algorithm in WEKA.
- **Instance-based Learning algorithms:** these algorithms use specific instances to achieve the classification predictions. The algorithms used are k-nearest neighbour one (k-NN) [54], where the classification is based on k nearest neighbours of a new instance (Ibk in WEKA) and K* [55], an instance-based classifier that uses an entropy-based distance function to classify data (kStar in WEKA).

It is important to remark here that other classification techniques are not reported in this study due to the poor performance achieved during our experiments.

IV. RESULTS AND DISCUSSION

Cross-validation was used in our experiments, to overcome the problem of overfitting and to make the predictions more general. In detail we have made reference to a 10-fold cross-validation, dividing the training set into k=10 smaller sets. For each of the k folds, a model is trained using k-1 of the folds

as the training data, while the resulting model is validated on the remaining part of the data.

The performance of the selected machine learning classification techniques was evaluated in terms of accuracy, sensibility, specificity and ROC area by using the following measurements:

- True Positive (TP): the voice sample is pathological and the algorithm recognizes this;
- True Negative (TN): the voice sample is healthy and the algorithm recognizes this;
- False Positive (FP): the voice sample is healthy but the algorithm recognizes it as pathological;
- False Negative (FN): the voice sample is pathological but the algorithm recognizes it as healthy.

The accuracy, that is the percentage of correctly classified instances, is defined as:

$$Accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)} \quad (3)$$

while the sensitivity and the specificity, that represent respectively the test's ability to detect positive results or the identification of negative results, are defined as:

$$Sensitivity = \frac{TP}{(TP + FN)} \quad (4)$$

$$Specificity = \frac{TN}{(TN + FP)} \quad (5)$$

The ROC area is a measure of the goodness of a classification algorithm evaluated by plotting a curve representing the sensitivity versus the complementarity of specificity (1-specificity) and measuring the area under this curve (AUC). The AUC can be interpreted as the average value of sensitivity for all the possible values of specificity. The maximum (AUC=1) means that the algorithm is perfect in the classification between diseased and non-diseased voices. On the other hand, AUC=0 means that the algorithm incorrectly classifies all subjects with diseases as negative and all healthy subjects as pathological [56].

A. Features selection

Attribute selection is an important task that allows the improvement of the dataset analysis to identify the redundant and/or irrelevant features to optimize memory space and time machine computing speed. For these reasons, in our study, we have chosen to test the machine learning classification techniques over the overall database, and, additionally, over three different subset of the database chosen by selecting some of the calculated features applying the following features selection methods:

- **InfoGainAttributeEval** algorithm [57]: this calculates the information gain for each feature. The results can vary from 0 (no information) to 1 (maximum information). The information gain obtained for each considered feature is shown in Figure 2. The best value is achieved by the age, followed by two MFCC coefficients, the HNR, jitter, the second derivative, and others. For our experimental

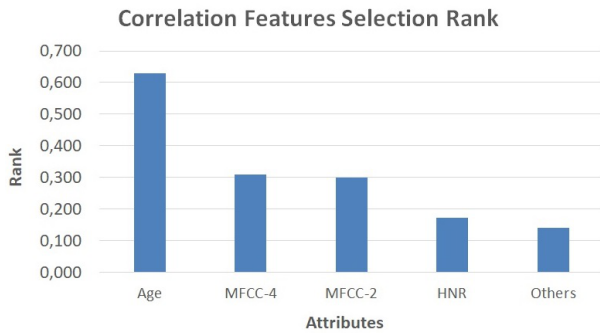


Fig. 3. Correlation rank obtained for each feature.

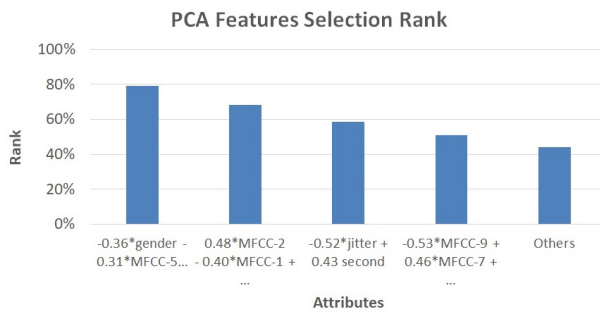


Fig. 4. PCA rank obtained in our study.

tests, we excluded all features that did not produce an information gain, this is equal to 0, while those features with an information gain greater than 0 were considered.

- *Correlation* method [58]: this assesses the predictive ability of each attribute, giving us the possibility of preferring sets of attributes that are highly correlated with the class. We have used 0.15 as our cut-off for relevant attributes, and the remaining attributes have been removed in accordance with the Figure 3.
- *Principal Components Analysis (PCA)* method [59]: Similarly to [60] we used the PCA method to select the most significant parameters. We selected the principal components which have obtained at least 50% of the ranking. As shown in Figure 4, we obtained four new parameters that are combination of several features.

The choice of these features selection methods was suggested by the wide use of such techniques in machine learning classifications to improve the overall quality of the patterns and/or the time required for the actual mining.

B. Classification Performance

Concerning the SMO algorithm, it is important to specify that we have conducted a series of tests in which we varied the values of the q exponent and γ parameter respectively in polynomial and RBF forms of the kernel expression. We have evaluated the accuracy, sensitivity, specificity and ROC area performance over the overall dataset and also over the three subset obtained from the features selection methods, indicated in Subsection IV-A. The results are reported in Tables III and IV.

The numerical results show that the best performance was, generally, obtained with a RBF kernel characterized by a γ -value equal to 1 over the database containing all parameters. For this value the SMO algorithm classifies with an accuracy of about 85.77% the presence of voice disorders in the analysed voices and a sensitivity of 87.59%, evaluating correctly a great percentage of signals, although the best results of sensitivity were obtained using a γ -value equal to 0.01 (about 91%). While the best performances obtained with a polynomial kernel were achieved with a q -value equal to 2 when the database containing all parameters. In this case, the classification accuracy was about 85.33% and the sensitivity and specificity were equal to, respectively, 87.88% and 82.77%.

Considering only the parameters selected with the *InfoGainAttribute Eval* method, the best performances were obtained, in the case of polynomial kernel, with a q -value equal to 3.5. The accuracy, sensitivity, specificity and ROC area were, respectively, equal to 83.50%, 83.94%, 83.07% and 0.835. Instead, considering the RBF kernel, the best results were achieved with a γ -value equal to 4. In this case the classification accuracy was about 84%.

Also, considering only parameters selected with the *Correlation* method, the γ -value equal to 4 was useful to achieve the best accuracy, although lower than previous case (82.99%). While with a polynomial kernel, the best performances were obtained with a q -value equal to 0.5. The accuracy is 82.34% while the sensitivity, specificity and ROC area were, respectively, 86.42%, 78.25% and 0.823%.

Finally, when we considered only parameters selected with *PCA* method, the best accuracy was achieved with the RBF kernel with a γ -value equal to 4, that is about 72%. While, with polynomial kernel, the q -value equal to 3 allowed to obtain the best classification accuracy, that is 70.22%.

Comparing the results obtained with several machine learning algorithms, shown in Table V, the best accuracy (85.77%) in the voice pathology detection was achieved using the SMO technique with a polynomial kernel characterized by a γ -value equal to 1 and using all parameters for the classification. Result confirmed observing the successive experimental tests where we have selected opportune features with two of the three features selection methods. Considering, in fact, only the features selected with the *InfoGainAttribute Eval* method and *PCA* one, the highest classification accuracy values were obtained using the SMO technique (respectively equal to 84.16% and 71.75%). Meanwhile, when we considered the features selected with *Correlation method*, the best accuracy to discriminate between pathological and healthy voices was achieved with the Decision Tree Algorithm equal to about 84%.

V. CONCLUSIONS

In recent years, the use of mobile multimedia services and applications in healthcare sector has been increasing significantly. Mobile health applications allow people to access medical information and data of interest at any time and anywhere, useful for the monitoring and detection of

TABLE III
CLASSIFICATION RESULTS OBTAINED BY THE SVM ALGORITHM FOR SEVERAL q VALUES.

<i>Considering all parameters</i>									
q	0.01	0.5	1	1.5	2	2.5	3	3.5	4
Accuracy (%)	49.64	82.77	83.14	84.23	85.33	85.26	85.11	85.04	82.48
Sensitivity (%)	49.64	87.59	87.30	87.88	87.88	86.42	85.69	84.53	82.92
Specificity (%)	49.64	77.96	78.98	80.58	82.77	84.09	84.53	85.55	82.04
ROC Area	0.496	0.828	0.831	0.842	0.853	0.853	0.851	0.85	0.825
<i>Considering only the parameters selected with the InfoGainAttribute Eval features selection</i>									
q	0.01	0.5	1	1.5	2	2.5	3	3.5	4
Accuracy (%)	50.29	81.97	81.82	81.90	82.70	82.48	83.28	83.50	83.50
Sensitivity (%)	50.95	86.86	86.13	85.84	85.40	83.50	83.65	83.94	84.23
Specificity (%)	49.64	77.08	77.52	77.96	80.00	81.46	82.92	83.07	82.77
ROC Area	0.503	0.82	0.818	0.819	0.827	0.825	0.833	0.835	0.835
<i>Considering only the parameters selected with the Correlation method features selection</i>									
q	0.01	0.5	1	1.5	2	2.5	3	3.5	4
Accuracy (%)	72.12	82.34	82.04	81.82	81.53	81.82	81.53	81.17	80.58
Sensitivity (%)	94.74	86.42	86.42	86.72	86.57	87.01	87.15	87.45	87.15
Specificity (%)	49.49	78.25	77.66	76.93	76.50	76.64	75.91	74.89	74.01
ROC Area	0.721	0.823	0.82	0.818	0.815	0.818	0.815	0.812	0.806
<i>Considering only the parameters selected with the PCA features selection</i>									
q	0.01	0.5	1	1.5	2	2.5	3	3.5	4
Accuracy (%)	49.64	69.49	69.78	69.78	70.00	70.00	70.22	70.00	70.00
Sensitivity (%)	49.64	83.80	81.46	79.85	78.10	76.64	75.62	73.72	72.41
Specificity (%)	49.64	55.18	58.10	59.71	61.90	63.36	64.82	66.28	67.59
ROC Area	0.496	0.695	0.698	0.698	0.7	0.7	0.702	0.7	0.7

TABLE IV
CLASSIFICATION RESULTS OBTAINED BY THE SVM ALGORITHM FOR SEVERAL γ VALUES.

<i>Considering all parameters</i>									
γ	0.01	0.5	1	1.5	2	2.5	3	3.5	4
Accuracy (%)	79.05	84.96	85.77	85.33	85.04	85.26	84.67	84.67	84.82
Sensitivity (%)	91.53	88.18	87.59	87.45	87.01	86.72	85.69	85.69	85.40
Specificity (%)	66.57	81.75	83.94	83.21	83.07	83.80	83.65	83.65	84.23
ROC Area	0.791	0.85	0.858	0.853	0.85	0.853	0.847	0.847	0.848
<i>Considering only the parameters selected with the InfoGainAttribute Eval features selection</i>									
γ	0.01	0.5	1	1.5	2	2.5	3	3.5	4
Accuracy (%)	77.88	82.55	82.70	82.77	83.50	83.65	83.72	83.87	84.16
Sensitivity (%)	91.39	84.82	83.80	82.77	84.23	84.38	84.38	84.96	85.26
Specificity (%)	64.38	80.29	81.61	82.77	82.77	82.92	83.07	82.77	83.07
ROC Area	0.779	0.826	0.827	0.828	0.835	0.836	0.837	0.839	0.842
<i>Considering only the parameters selected with the Correlation method features selection</i>									
γ	0.01	0.5	1	1.5	2	2.5	3	3.5	4
Accuracy (%)	78.18	82.04	82.77	82.63	82.55	82.55	82.77	82.85	82.99
Sensitivity (%)	90.95	85.26	83.36	82.48	82.63	82.63	82.77	82.92	83.07
Specificity (%)	65.40	78.83	82.19	82.77	82.48	82.48	82.77	82.77	82.92
ROC Area	0.782	0.82	0.828	0.826	0.847	0.847	0.826	0.828	0.83
<i>Considering only the parameters selected with the PCA features selection</i>									
γ	0.01	0.5	1	1.5	2	2.5	3	3.5	4
Accuracy (%)	53.65	70.07	70.73	71.09	71.39	71.02	71.02	71.68	71.75
Sensitivity (%)	57.96	81.75	81.61	82.63	82.77	82.63	82.19	82.34	82.34
Specificity (%)	49.34	58.39	59.85	59.56	60.00	59.42	59.85	61.02	61.17
ROC Area	0.536	0.701	0.707	0.711	0.714	0.71	0.71	0.717	0.718

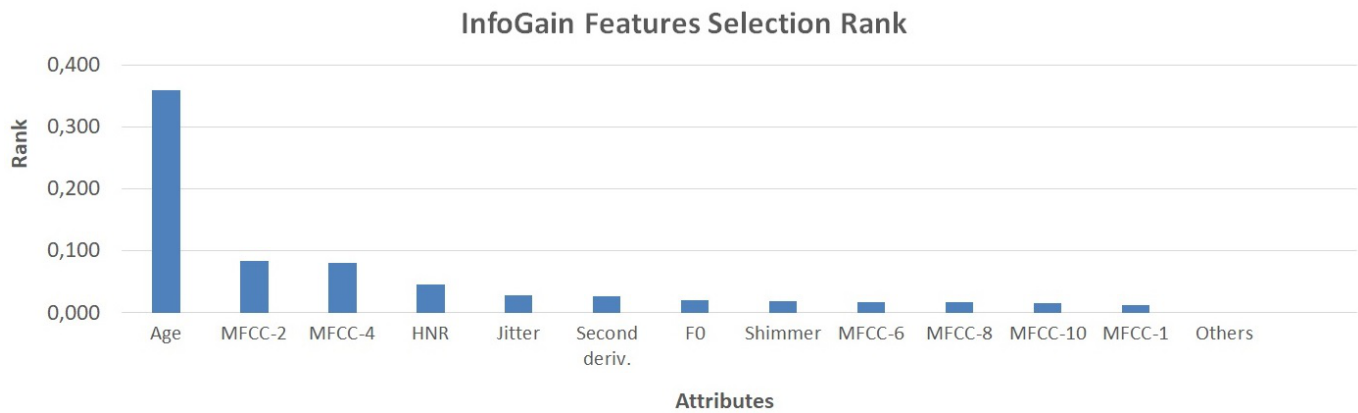


Fig. 2. Information gain estimated for each feature.

TABLE V
CLASSIFICATION RESULTS FOR DIFFERENT ALGORITHMS.

<i>Considering all parameters</i>						
	<i>SMO</i>	<i>DT</i>	<i>BC</i>	<i>LMT</i>	<i>Ibk</i>	<i>Kstar</i>
Accuracy (%)	85.77	82.04	78.61	82.48	76.28	49.71
Sensitivity (%)	87.59	81.02	79.85	86.28	77.37	64.09
Specificity (%)	83.94	83.07	77.37	78.69	75.18	35.33
ROC Area	0.858	0.832	0.863	0.885	0.763	0.524
<i>Considering only the parameters selected with the InfoGainAttribute Eval features selection</i>						
	<i>SMO</i>	<i>DT</i>	<i>BC</i>	<i>LMT</i>	<i>Ibk</i>	<i>Kstar</i>
Accuracy (%)	84.16	82.48	78.61	82.12	76.20	50.22
Sensitivity (%)	85.26	80.58	79.85	86.13	77.23	64.82
Specificity (%)	83.07	84.38	77.37	78.10	75.18	35.62
ROC Area	0.842	0.830	0.863	0.883	0.762	0.535
<i>Considering only the parameters selected with the Correlation method features selection</i>						
	<i>SMO</i>	<i>DT</i>	<i>BC</i>	<i>LMT</i>	<i>Ibk</i>	<i>Kstar</i>
Accuracy (%)	82.99	83.58	80.58	82.26	77.45	81.82
Sensitivity (%)	83.07	83.80	81.61	86.72	78.39	84.23
Specificity (%)	82.92	83.36	79.56	77.81	76.50	79.42
ROC Area	0.83	0.857	0.873	0.884	0.774	0.873
<i>Considering only the parameters selected with the PCA features selection</i>						
	<i>SMO</i>	<i>DT</i>	<i>BC</i>	<i>LMT</i>	<i>Ibk</i>	<i>Kstar</i>
Accuracy (%)	71.75	69.20	69.12	69.12	64.53	67.66
Sensitivity (%)	82.34	62.92	65.99	74.45	64.67	72.12
Specificity (%)	61.17	75.47	72.26	63.80	64.38	63.21
ROC Area	0.718	0.737	0.745	0.753	0.645	0.737

specific diseases, such as dysphonia, a voice disorder often underestimated that affects a great percentage of people.

Research on mobile automatic systems to estimate voice disorders has received considerable attention in the last few years due to its objectivity and non-invasive nature. Machine learning techniques can be a valid support to investigate new approaches to signal processing in an easy and fast way that can be implemented in an m-health solution. This study compares the performance of different voice pathology identification methods, taking into account the main machine learning techniques. Several techniques are applied such as the Support Vector Machine, Decision Tree, Bayesian Clas-

sification, Logistic Model Tree and Instance-based Learning algorithms. Moreover, in this work we focus on identifying appropriate voice signal features by using the comparative study of different classifiers. All analyses are performed on a wide dataset of 1370 voices selected from the Saarbruecken Voice Database.

The tests have been carried out over the overall dataset and over three different subset where we only have considered the selected features by three specific features selection methods. The results have shown that the best accuracy in voice pathology detection is achieved using the Support Vector Machine algorithm. This technique classifies a voice as pathological or healthy with an accuracy equal to about 86% using all parameters. Result confirmed observing the experimental tests in which the *InfoGainAttribute Eval* method and *PCA* one have been applied (accuracy values were, respectively, equal to 84.16% and 71.75%). Meanwhile, when we only considered parameters selected with the *Correlation* method, the best accuracy was obtained with the Decision Tree technique.

Although the accuracy values are smaller than the values obtained in other studies in literature, it is necessary to highlight that all these studies are performed on very limited and, often, non-accessible datasets. To enhance the classification rate obtained we are interested in improving the classification phase by developing a hybrid system using a combination of several machine learning techniques. In future work, we want to integrate this hybrid classifier in an m-health system, such as the solution proposed in [23], able to detect the presence or not of a voice disorders, useful to monitoring and treatment of patients suffering from these pathologies.

REFERENCES

- [1] B. M. Silva, J. J. Rodrigues, I. de la Torre Díez, M. López-Coronado, and K. Saleem, "Mobile-health: A review of current state in 2015," *Journal of biomedical informatics*, vol. 56, pp. 265–272, 2015.
- [2] S. Naddeo, L. Verde, M. Forastiere, G. De Pietro, and G. Sannino, "A real-time m-health monitoring system: An integrated solution combining the use of several wearable sensors and mobile devices." in *HEALTH-INF*, 2017, pp. 545–552.
- [3] M. S. Hossain and G. Muhammad, "Cloud-assisted industrial internet of things (iiot)-enabled framework for health monitoring," *Computer Networks*, vol. 101, pp. 192–202, 2016.

- [4] G. Sannino, I. De Falco, and G. De Pietro, "A supervised approach to automatically extract a set of rules to support fall detection in a mhealth system," *Applied Soft Computing*, vol. 34, pp. 205–216, 2015.
- [5] J. Mohammed, C.-H. Lung, A. Ocleanu, A. Thakral, C. Jones, and A. Adler, "Internet of things: Remote patient monitoring using web services and cloud computing," in *Internet of Things (iThings), 2014 IEEE International Conference on, and Green Computing and Communications (GreenCom), IEEE and Cyber, Physical and Social Computing (CPSCom), IEEE*. IEEE, 2014, pp. 256–263.
- [6] R. Gravina and G. Fortino, "Automatic methods for the detection of accelerative cardiac defense response," *IEEE Transactions on Affective Computing*, vol. 7, no. 3, pp. 286–298, 2016.
- [7] S. Iyengar, F. T. Bonda, R. Gravina, A. Guerrieri, G. Fortino, and A. Sangiovanni-Vincentelli, "A framework for creating healthcare monitoring applications using wireless body sensor networks," in *Proceedings of the ICST 3rd international conference on Body area networks*. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2008, p. 8.
- [8] R. H. G. Martins, H. A. do Amaral, E. L. M. Tavares, M. G. Martins, T. M. Gonçalves, and N. H. Dias, "Voice disorders: etiology and diagnosis," *Journal of Voice*, vol. 30, no. 6, pp. 761–e1, 2016.
- [9] A. R. Maccarini and E. Lucchini, "La valutazione soggettiva ed oggettiva della disfonia. il protocollo sifel," *ACTA PHONIATRICA LATINA*, vol. 24, no. 1/2, pp. 13–42, 2002.
- [10] J. I. Godino-Llorente, P. Gómez-Vilda, N. Sáenz-Lechón, M. Blanco-Velasco, F. Cruz-Roldán, and M. A. Ferrer-Ballester, "Support vector machines applied to the detection of voice disorders," *Lecture notes in computer science*, vol. 3817, p. 219, 2005.
- [11] N. Souissi and A. Cherif, "Dimensionality reduction for voice disorders identification system based on mel frequency cepstral coefficients and support vector machine," in *Modelling, Identification and Control (ICMIC), 2015 7th International Conference on*. IEEE, 2015, pp. 1–6.
- [12] H. Cordeiro, J. Fonseca, I. Guimarães, and C. Meneses, "Voice pathologies identification speech signals, features and classifiers evaluation," in *Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA), 2015*. IEEE, 2015, pp. 81–86.
- [13] F. AMARA, M. FEZARI, and H. BOUROUBA, "An improved gmm-svm system based on distance metric for voice pathology detection," *Appl. Math*, vol. 10, no. 3, pp. 1061–1070, 2016.
- [14] M. El Emary, M. Fezari, and F. Amara, "Towards developing a voice pathologies detection system," *Journal of Communications Technology & Electronics*, vol. 59, no. 11, p. 1280, 2014.
- [15] E. S. Fonseca, R. C. Guido, P. R. Scalassara, C. D. Maciel, and J. C. Pereira, "Wavelet time-frequency analysis and least squares support vector machines for the identification of voice disorders," *Computers in Biology and Medicine*, vol. 37, no. 4, pp. 571–578, 2007.
- [16] J. Wang and C. Jo, "Vocal folds disorder detection using pattern recognition methods," in *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*. IEEE, 2007, pp. 3253–3256.
- [17] R. Ritchings, M. McGillion, and C. Moore, "Pathological voice quality assessment using artificial neural networks," *Medical engineering & physics*, vol. 24, no. 7, pp. 561–564, 2002.
- [18] B. Boyanov and S. Hadjitodorov, "Acoustic analysis of pathological voices. a voice analysis system for the screening of laryngeal diseases," *IEEE Engineering in Medicine and Biology Magazine*, vol. 16, no. 4, pp. 74–82, 1997.
- [19] S. Jothilakshmi, "Automatic system to detect the type of voice pathology," *Applied Soft Computing*, vol. 21, pp. 244–249, 2014.
- [20] N. Saenz-Lechon, J. I. Godino-Llorente, V. Osmá-Ruiz, and P. Gómez-Vilda, "Methodological issues in the development of automatic systems for voice pathology detection," *Biomedical Signal Processing and Control*, vol. 1, no. 2, pp. 120–128, 2006.
- [21] M. S. Hossain, G. Muhammad, M. F. Alhamid, B. Song, and K. Al-Mutib, "Audio-visual emotion recognition using big data towards 5g," *Mobile Networks and Applications*, vol. 21, no. 5, pp. 753–763, 2016.
- [22] M. S. Hossain, "Patient state recognition system for healthcare using speech and facial expressions," *Journal of medical systems*, vol. 40, no. 12, p. 272, 2016.
- [23] L. Verde, G. De Pietro, P. Veltri, and G. Sannino, "An m-health system for the estimation of voice disorders," in *Multimedia & Expo Workshops (ICMEW), 2015 IEEE International Conference on*. IEEE, 2015, pp. 1–6.
- [24] M. Mat Baki, G. Wood, M. Alston, P. Ratcliffe, G. Sandhu, J. Rubin, and M. Birchall, "Reliability of operavox against multidimensional voice program (mdvp)," *Clinical Otolaryngology*, vol. 40, no. 1, pp. 22–28, 2015.
- [25] E. Van Leer, R. C. Pfister, and X. Zhou, "An ios-based cepstral peak prominence application: feasibility for patient practice of resonant voice," *Journal of Voice*, vol. 31, no. 1, pp. 131–e9, 2017.
- [26] R. Gravina, P. Alinia, H. Ghasemzadeh, and G. Fortino, "Multi-sensor fusion in body sensor networks: State-of-the-art and research challenges," *Information Fusion*, vol. 35, pp. 68–80, 2017.
- [27] M. Chen, P. Zhou, and G. Fortino, "Emotion communication system," *IEEE Access*, vol. 5, pp. 326–337, 2017.
- [28] K. Elemetrics, "Kay elemetrics corp. disordered voice database," *Model*, vol. 4337, 1994.
- [29] W. Barry and M. Pützer, "Saarbrücken voice database," *Institute of Phonetics, Universität des Saarlandes*, <http://www.stimmdatenbank.coli.uni-saarland.de>, 2007.
- [30] I. H. Witten, E. Frank, M. A. Hall, and C. J. Pal, *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann, 2016.
- [31] M. Pützer and J. Koreman, "A german database of patterns of pathological vocal fold vibration," *Phonus*, vol. 3, pp. 143–153, 1997.
- [32] D. Martínez, E. Lleida, A. Ortega, A. Miguel, and J. Villalba, "Voice pathology detection on the saarbrücken voice database with calibration and fusion of scores using multifocal toolkit," in *Advances in Speech and Language Technologies for Iberian Languages*. Springer, 2012, pp. 99–109.
- [33] A. Tsanas, "Acoustic analysis toolkit for biomedical speech signal processing: concepts and algorithms," *Models and analysis of vocal emissions for biomedical applications*, vol. 2, pp. 37–40, 2013.
- [34] L. W. Lopes, L. B. Simões, J. D. da Silva, D. da Silva Evangelista, A. C. da Nóbrega, P. O. C. Silva, V. J. D. Vieira *et al.*, "Accuracy of acoustic analysis measurements in the evaluation of patients with different laryngeal diagnoses," *Journal of Voice*, vol. 31, no. 3, pp. 382–e15, 2017.
- [35] C. Manfredi, M. D'Aniello, and P. Brusciaglioni, "Acoustic measure of noise energy in vocal folds operated patients," in *Signal Processing Conference (EUSIPCO 1998), 9th European*. IEEE, 1998, pp. 1–4.
- [36] S. Y. Lowell, R. H. Colton, R. T. Kelley, and Y. C. Hahn, "Spectral- and cepstral-based measures during continuous speech: capacity to distinguish dysphonia and consistency within a speaker," *Journal of Voice*, vol. 25, no. 5, pp. e223–e232, 2011.
- [37] G. Schlotthauer, M. E. Torres, and H. L. Rufner, "Pathological voice analysis and classification based on empirical mode decomposition," in *Development of multimodal interfaces: active listening and synchrony*. Springer, 2010, pp. 364–381.
- [38] D. Talkin, "A robust algorithm for pitch tracking (rapt)," *Speech coding and synthesis*, vol. 495, p. 518, 1995.
- [39] P. A. Naylor, A. Kounoudes, J. Gudnason, and M. Brookes, "Estimation of glottal closure instants in voiced speech using the dyspa algorithm," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 15, no. 1, pp. 34–43, 2007.
- [40] H. Kawahara, A. d. Cheveigné, H. Banno, T. Takahashi, and T. Irino, "Nearly defect-free f0 trajectory extraction for expressive speech modifications based on straight," in *Ninth European Conference on Speech Communication and Technology*, 2005.
- [41] L. Tan and M. Karnjanadecha, "Pitch detection algorithm: autocorrelation method and amdf," in *Proceedings of the 3rd international symposium on communications and information technology*, vol. 2, 2003, pp. 551–556.
- [42] L. Verde, G. De Pietro, and G. Sannino, "A methodology for voice classification based on the personalized fundamental frequency estimation," *Biomedical Signal Processing and Control*, vol. 42, pp. 134–144, 2018.
- [43] A. De Cheveigné and H. Kawahara, "Yin, a fundamental frequency estimator for speech and music," *The Journal of the Acoustical Society of America*, vol. 111, no. 4, pp. 1917–1930, 2002.
- [44] M. Farrús, J. Hernandez, and P. Ejarque, "Jitter and shimmer measurements for speaker recognition," in *Eighth Annual Conference of the International Speech Communication Association*, 2007.
- [45] F. Severin, B. Bozkurt, and T. Dutoit, "Hnr extraction in voiced speech, oriented towards voice quality analysis," in *Proc. EUSIPCO'05*, 2005.
- [46] "VOICEBOX: speech processing toolbox for matlab," <http://www.ee.ic.ac.uk/hp/staff/dmb/voicebox/voicebox.html>, accessed: 2018-01-23.
- [47] T. Marciniak, R. Weychan, S. Drgas, A. Dabrowski, and A. Krzykowska, "Speaker recognition based on short polish sequences," in *Signal Processing Algorithms, Architectures, Arrangements, and Applications Conference Proceedings (SPA), 2010*. IEEE, 2010, pp. 95–98.
- [48] B. Schölkopf, C. J. Burges, and A. J. Smola, *Advances in kernel methods: support vector learning*. MIT press, 1999.
- [49] V. N. Vapnik, "An overview of statistical learning theory," *IEEE transactions on neural networks*, vol. 10, no. 5, pp. 988–999, 1999.

- [50] E. Alpaydin, *Introduction to machine learning*. MIT press, 2014.
- [51] S. L. Salzberg, "C4. 5: Programs for machine learning by j. ross quinlan. morgan kaufmann publishers, inc., 1993," *Machine Learning*, vol. 16, no. 3, pp. 235–240, 1994.
- [52] G. H. John and P. Langley, "Estimating continuous distributions in bayesian classifiers," in *Proceedings of the Eleventh conference on Uncertainty in artificial intelligence*. Morgan Kaufmann Publishers Inc., 1995, pp. 338–345.
- [53] N. Landwehr, M. Hall, and E. Frank, "Logistic model trees," *Machine learning*, vol. 59, no. 1-2, pp. 161–205, 2005.
- [54] D. W. Aha, D. Kibler, and M. K. Albert, "Instance-based learning algorithms," *Machine learning*, vol. 6, no. 1, pp. 37–66, 1991.
- [55] J. G. Cleary and L. E. Trigg, "K*: An instance-based learner using an entropic distance measure," in *Proceedings of the 12th International Conference on Machine learning*, vol. 5, 1995, pp. 108–114.
- [56] K. Hajian-Tilaki, "Receiver operating characteristic (roc) curve analysis for medical diagnostic test evaluation," *Caspian journal of internal medicine*, vol. 4, no. 2, p. 627, 2013.
- [57] I. H. Witten and E. F. D. Mining, "Practical machine learning tools and techniques morgan kaufmann," 2005.
- [58] "CorrelationAttributeEval," <http://weka.sourceforge.net/doc.dev/weka/attributeSelection/CorrelationAttributeEval.html>, retrieved Januaray 29, 2018.
- [59] I. Jolliffe, "Principal component analysis," *Department of Mathematical Sciences, Kings College*, 2002.
- [60] M. K. Arjmandi, M. Pooyan, H. Mohammadnejad, and M. Vali, "Voice disorders identification based on different feature reduction methodologies and support vector machine," in *Electrical Engineering (ICEE), 2010 18th Iranian Conference on*. IEEE, 2010, pp. 45–49.



Giovanna Sannino received a Master's degree in Telecommunications Engineering (Cum Laude) from the University of Naples Parthenope in April 2011 and a Ph.D. degree in Information Engineering at the University of Naples "Parthenope" in April 2015. In 2009, she earned a Master degree, named "European Master on Critical Networked Systems", and a Bachelor's degree in Computer Engineering in 2008 at the University of Naples Federico II. Currently, she is a researcher at Institute for the High-Performance Computing and Networking (ICAR-CNR) and Adjunct Professor of Informatics at the University of Naples Federico II. Her research activities are focused on the areas of mobile Health, Pervasive Computing, Pattern Recognition, Signal Processing, and Artificial Intelligence for Healthcare. In addition, she is an IEEE 11073 Personal Health Device Working Group Member.



Laura Verde received a Bachelor's degree in Biomedical Engineering in 2009 and a Master's degree in 2011 at "Federico II" University in Naples (Italy). She is currently a PhD student in Information Engineering at the University of Naples Parthenope and collaborates with the Institute of High-Performance Computing and Networking (ICAR) of the National Research Council of Italy (CNR), Naples, Italy. Her research interests include Biomedical Signal Processing, Voice Signal Analysis, m-Health Systems and Artificial Intelligence for

Healthcare.



Giuseppe De Pietro is Director of the Institute for High Performance Computing and Networking - CNR, and Adjunct Professor at Temple University College of Science and Technology - Philadelphia. His current research interests are focused on cognitive computing, clinical decision support systems, software architectures for e-health. He has been actively involved in many European and National projects, with industrial co-operations. He is author of over 200 scientific papers published in international journals and conferences, and he is involved

in many program committees and journal editorial boards. He is an IEEE and KES International member.