Arabian Journal for Science and Engineering
https://doi.org/10.1007/s13369-020-04357-1

RESEARCH ARTICLE-ELECTRICAL ENGINEERING

Simple Logistic Hybrid System Based on Greedy Stepwise Algorithm for Feature Analysis to Diagnose Parkinson's Disease According to Gender

Şule Yücelbaş[1]

Received: 19 October 2019 / Accepted: 19 January 2020

Abstract
The incidence of Parkinson's disease (PD) is higher in males than in females. This disease can be diagnosed based on gender through the automatic diagnostic system without visiting a specialist physician. For this purpose, the Simple Logistic hybrid system based on the greedy stepwise search algorithm (SLGS) is presented as a novel approach that involves feature analysis to diagnose PD by gender. Six feature groups were extracted from data of 252 subjects and were used to perform this analysis. When the SLGS hybrid system analysis begins, the effective features are first determined by the greedy stepwise search algorithm and diagnosis of PD was made according to the features that were determined using the Simple Logistic classifier. The performance results that were obtained from this system by gender were evaluated using many statistical metrics. The accuracy ratios were 88.71% and 87.15%, with the minimum number of features for males and females, respectively. Additionally, ROC and PRC area values for both genders approached the 0.9 band, showing that the class separation of patients and healthy individuals was nearly perfect. The performance results and perspective of this study were compared with the single published study that used the same data, and the SLGS hybrid system showed higher performance rates than the other study. In addition, the number of the subjects was higher in this study than in other studies and the SLGS system has not been used before in the literature in this field.

Keywords Gender analysis · Greedy stepwise search · Parkinson's disease · Speech signal analysis

1 Introduction

Cells in certain parts of brain are responsible for producing the chemical dopamine [1]. Dopamine is a substance that allows people to act smoothly and harmoniously [1]. The deficiency of this substance is the most important factor in the development of Parkinson's disease (PD), which is a neurological disorder [1]. As a result of the loss of 60–80% of the cells producing dopamine during advanced age, dopamine levels become insufficient. Thus, motor disorders, are the most prominent symptoms of PD, begin to occur [1]. The first symptoms in people with this disease occur between the ages of 40 and 70 years old; nevertheless, most of these symptoms occur in individuals in their 60s [1]. Today, one of every 100 people over 65 years old is considered to have PD [1]. There are many symptoms in people with this disease. However, the initial symptoms allowing a PD diagnosis are slow motion and tremors during rest [1, 2]. In addition to these symptoms, dullness in facial expressions, disordered contractions, posture, gait, speech and smell, as well as forgetfulness and restless leg syndrome may occur [1]. Foulds et al. [3] showed that it is possible to diagnose PD at an early stage with a simple blood test. In another study, researchers tried to detect early stages of PD using tomography images [4].

Nowadays, gender is known to play an important role in the incidence and treatment of many diseases. Every cell in the human body has a gender [5, 6]. Each cell includes two X chromosomes in females, or both X and Y chromosomes in males. This chromosomal difference affects the health, disease and treatment methods of individuals [5, 6]. The first important step in the development of individualized treatments is to determine the effects of gender difference on the relevant disease [5, 6]. For example, a heart attack in males

✉ Şule Yücelbaş
suleyucelbas@hakkari.edu.tr

[1] Electrical and Electronics Engineering Department, Hakkari University, 30000 Hakkari, Turkey

Published online: 01 February 2020

is mostly caused by plaque rupture; however, in females, the underlying cause is typically plaque erosion [5, 6]. Another example can be given for Alzheimer's. The incidence of this disease in females is two times higher than in males [5, 6]. PD is another disease in which gender is an important factor for diagnosis and treatment. PD is seen twice as much in men than in women [5–7]. The estrogen hormone is known to be one of the reasons this disease is less common in females [5–7]. It has been found that this hormone has a positive effect on dopamine producing cells, neurons and pathways in the brain [5–7]. In addition, disease symptoms occur in women around two years later than in men [5, 7, 8]. Women with PD are more affected by anxiety, sadness, depression and constipation, whereas in men with this disease, there are more daytime sleepiness, dribbling and other gender-related symptoms [5, 7, 8]. In summary, many diseases such as PD are affected by gender differences [5, 7, 8]. The severity and progression of PD are more common in men than women, and it should be emphasized that the molecular details have an important place in these disease characteristics [5, 7, 8]. In the treatment of this disease, it is thought that scientists will eventually develop hormonal-based treatment methods, which will lead to greater progress [5, 7, 8].

The effects of gender on PD, the determination of the symptoms of this disease according to gender and gender-based treatment methods have mostly been carried out in the field of medicine [6–18]. In the field of engineering, studies on PD have generally been based on automatic detection. For this purpose, voice or walking data recorded from PD patients and healthy individuals were mostly used. Speech disorder is one of the most important symptoms that emerges in most PD patients [19, 20]. Therefore, in most of the PD engineering studies, audio signals recorded in the hospital environment were used. These studies used features extracted from voice recordings in attempt to separate PD and healthy patients [20–25]. For example, an automatic diagnosis of PD was attempted by using a tele-diagnosis system designed in a study [21]. The features extracted from the signals of patient and healthy subjects were given to the support vector machine (SVM) classifier, and designed system was tested [21]. In the study, maximum performance was tried to be achieved with the least features [21]. In another study, the authors tried various classification systems for PD detection and achieved the best accuracy ratio with 96.4% at SVM [26]. In a different study, after the Mel-frequency cepstral coefficients (MFCCs) technique was applied to the speech signals of the patient and healthy persons, the values were given to the SVM classifier [27]. When the results were evaluated, it was seen that /*u*/ letter records were more distinctive than other sounds [27]. In addition, many other studies [28–49] in recent years have addressed the automatic diagnosis of PD based on the sound signal. In one of these studies, researchers used a bio-inspired algorithm named Modified Gray Wolf Optimization (MGWO) to detect early signs of PD, which was evaluated on various datasets such as voice, handwriting and speech, with an accuracy of 94% [28]. In another study using voice signals, 18 feature extraction methods and four machine learning algorithms were used and achieved a comparable accuracy rate [29]. Zhang designed a machine learning-based PD tele-diagnosis system for smartphones [30]. In [31], a combined classification algorithm was introduced and applied on speech samples for the same purpose, which improved the accuracy of classification results. In another study, MFCC and PLP features were extracted from vocal recordings from PD patients and healthy subjects by using Single Taper Smooth (STS) and Thomson Multitaper (TMT) windowing methods [32]; the performance of TMT was found to be superior to that of STS [32]. In [33], authors investigated the severity score of PD with six different regression methods [Linear, Stepwise, Lasso, Ridge regression, prediction using a neural network (NN) model and classification and regression trees (CART)]. Results revealed the lowest error values with the NN model (1.5) and with CART (1.8) [33]. In [34], Parisi et al. proposed a hybrid feature-driven algorithm [multilayer perceptron–Lagrangian support vector machine (MLP–LSVM)], and the distinction between PD patients and healthy subjects reached 100% accuracy. Another study used empirical wavelet transform (EWT) and empirical wavelet packet transform (EWPT) for detection of three different PD severities [35]. Three classifiers [*k*-nearest neighbor (KNN), probabilistic neural network (PNN) and extreme learning machine (ELM)] were also used, with classification results of more than 90% and 95% for EWT and EWPT, respectively [35]. In [36], Mamun et al. introduced a novel approach consisting of a cloud-based framework for diagnosis of PD. The proposed system achieved a classification accuracy of 96.6% [36]. In [37], a novel hybrid system was proposed for the detection of PD that includes three steps, namely feature processing, feature reduction/selection and classification [37]. In [38], Haq et al. proposed a machine learning-based recognition system wherein a support vector machine (SVM) was employed to detect PD, achieving an accuracy rate of 99%. In a similar study, extreme learning machine (ELM) and kernel ELM systems were investigated for automatic detection of PD, with an average accuracy rate of 95.97% [39]. A classification accuracy of 97.42% was obtained in another study whose primary aim was to analyze the capabilities of SVM in a SVM-based bacterial foraging optimization (BFO) hybrid system to detect PD [40]. In [41], Ali et al. [41]

introduced an artificial intelligence system based on feature selection to detect PD, which yielded an accuracy rate of 100%. Likewise, 100% accuracy was obtained in research wherein the authors aimed to diagnose PD using a neuro-fuzzy system with 16 hidden nodes [42]. In [43], Wu et al. presented a new probabilistic-based class-confusion information measure technique named interclass probability risk (ICPR) and obtained high overall performance results. In another study, an adaptive artificial bee colony (AABC)–kernel-based weighted ELM (KWELM) system was proposed to diagnose PD, reaching 98.97% classification accuracy [44]. In [45], multiple feature evaluation (MFE) and classification methods [Decision Tree, naïve Bayes (NB), NN, random forests (RF) and SVM] were implemented for diagnosis of PD. The main purpose of this study was to investigate the effect of the MFE method on classification results. When the effect of the proposed MFE method on the results was examined, it was found to improve classification by a minimum of 9.13% and a maximum 15.22% [45]. In [46], Lahmiri and Shmuel applied the feature ranking techniques with SVM for detecting PD and obtained the highest accuracy rate of 92.21%. In another study, the performance of different classification systems (linear discriminant analysis, KNN, NB, regression trees, radial basis function NN, SVM and Mahalanobis distance) in the diagnosis of PD were compared, revealing that the SVM classifier performs best [47]. In [48], Cai et al. introduced an integrated system named chaotic bacterial foraging optimization-reinforced fuzzy KNN (CBFO-FKNN). This proposed system was found to yield better results than the other five FKNN-based methods (BFO, particle swarm optimization, genetic algorithms, fruit fly optimization and firefly algorithm) [48]. In [49], a new approach, named optimized cuttlefish, which uses the traditional cuttlefish algorithm, was presented. In addition, decision tree and KNN systems were used as classifiers, with the best result obtained at approximately 94% [49].

It is possible to explain more of the studies in the literature in this way. However, most of the studies are similar to each other because they are practiced in almost the same way and for the same purpose. Only the methods used differ slightly. It is understood that the common points of the studies in the literature are data recording, data processing, feature extraction, feature selection and classification. Also, the results in literature showed that the feature evaluation/selection/reduction methods are very significant in improving the accuracy rate. However, in these engineering investigations, detailed analyses regarding the features extracted from speech or gait signals, according to gender, were not investigated. This deficiency in the literature has motivated the following study. In this study, a new hybrid system that has not been previously applied to any dataset (in this field) was proposed. The model, which is presented as a new approach, is the Simple Logistic hybrid system and is based on the greedy stepwise search algorithm (SLGS). The SLGS hybrid system was used for feature analysis to diagnose PD according to gender. The main innovation of this approach is the automatic analysis of effective features that are extracted from the speech signals for the diagnosis of PD, according to gender, with high accuracy. For the proposed system, the features extracted from the audio signals of 252 people previously recorded in the Department of Neurology at Cerrahpaşa Faculty of Medicine, Istanbul University [50] were taken from the UCI machine learning repository. Using the k-fold CV method, the data were separated as training tests and introduced to SLGS system. Firstly, effective features were determined, and then, a diagnostic process was performed with the proposed system according to the determined features. The performance results obtained from the SLGS system (by gender) were evaluated with statistical criteria that included true positive (TP) rate, false positive (FP) rate, F-score (F), Matthews correlation coefficient (MCC), precision, receiver operating characteristics (ROC) and precision-recall (PRC) areas. With the proposed SLGS system, the effect of the Greedy Stepwise search algorithm, which has not been previously used for PD diagnosis, was demonstrated on feature analysis. In addition, thanks to the created SLGS hybrid system, a maximum success rate was achieved with a minimum number of features. Because there are no other studies that create and test an engineering-based diagnosis method for PD that is according to gender, this study is important.

2 Materials and Methods

PD is a progressive disorder of the nervous system that affects human movements. This disease sometimes manifests itself in small tremors that appear only in one hand of patient and progresses very slowly. In this disease, many symptoms can be seen in the early period and one of them is the changes in the pronunciation. In the person with this disease, the speech softens and deteriorates. Mumble in speaking, hesitation before speaking, very fast or very slow conversation may also occur [1]. The diagnosis of PD is facilitated by some of the features extracted from the speech signals of the patient and healthy subjects. For example, Fig. 1 shows the variation in the first 15 of the features extracted by applying the MFCCs method to audio signals taken from a healthy and PD subject in this study. When Fig. 1 is examined, it is

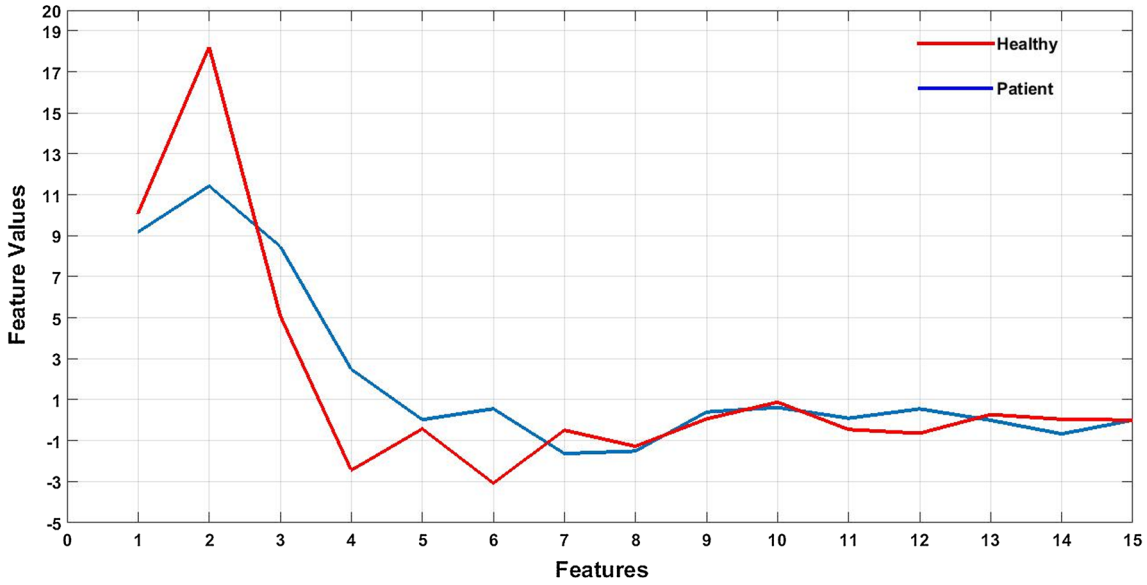

Fig. 1 Sample variation of speech signal features of healthy and PD subjects

understood that many of these features are distinctly separated from each other.

2.1 Dataset and Features

The dataset used in this study was taken from UCI machine learning repository [50]. The features taken by different methods from the speech signals of 252 subjects in the Department of Neurology at Cerrahpaşa Faculty of Medicine, Istanbul University, ranging in age between 33 and 87 were used in the dataset [50]. Of these subjects, 188 were Parkinson and 64 were healthy. The gender group containing the patient data consisted of 107 male and 81 female patients [50]. Besides, healthy data were collected from 23 male and 41 female subjects. Under physician supervision, each subject was made to repeat /*a*/ letter three times, the data were recorded and the microphone was arranged as 44.1 kHz during this process [50]. Six different feature group data previously obtained from the audio signals of 252 subjects in total were used for performed processes [50].

- The first group was baseline feature group (21 properties in total). It has the most common features related to audio signals in PD studies in the literature, consisted of jitter, shimmer, fundamental frequency parameters, harmonicity parameters, recurrence period density entropy (RPDE), detrended fluctuation analysis (DFA) and pitch period entropy (PPE) features [23, 24, 26, 50–52].
- The second group was time frequency feature group (11 properties in total), which occurred of intensity parameters, formant frequencies and bandwidth features.
- MFCCs were the third group, and these features are often used for voice recognition, biomedical sound analysis and PD diagnosis [52–55]. The MFCC method was used by [56] and [57] for the first time in the literature for automatic sound analysis. In the dataset used in this study, there were 84 properties obtained by applying the MFCCs method to audio signals, previously [50].
- The fourth feature group was the vocal fold attributes (22 properties in total), which consisted of glottis quotient (GQ), glottal to noise excitation (GNE), vocal fold excitation ratio (VFER) and empirical mode decomposition (EMD).
- Wavelet Transform-based Features was the fifth feature group (182 properties in total).
- The last group highlighted consisted of TQWT features. The TQWT method is a new discrete wavelet transform form consisting of three basic parameters, Q (Q-factor), j (the number of levels) and r (redundancy) [58]. The low and high frequency values of the signal can be separated, and also band-pass filters with different Q factors are generated using this method. The Q, found by dividing the bandwidth of the center frequency of a band-pass filter, gives this method its name. This factor, which is adjusted according to the oscillation of the processed signal, creates a nonlinear separation. If the frequency

distribution of the signal is examined, it is seen that the frequency spectrum of the sudden changing finite signals for the low Q-factor is wider. For the high Q-factor, the frequency spectra of the oscillation signals are more localized, so the Q-factor indicates the oscillation of the processed signals. j is defined as level number which will have $j+1$ sub-bands after high-pass filter and last low-pass filter outputs. The frequency of the band-pass filters is determined by the last parameter r, and as a result of this parameter, TQWT starts to resemble the continuous wavelet transform [58]. In this study; Q, r and j parameters were determined as 2, 4 and 35, respectively, in order to obtain the TQWT feature group [50]. In order to create this feature group, 12 different attributes were extracted from 36 sub-bands obtained using TQWT and 432 parameters were recorded (12×36) by creators in [50]. These features include energy, Shannon entropy, log energy entropy, Teager–Kaiser energy operator (TKEO) mean, TKEO standard deviation (std), median value, mean value, std value, minimum (min) value, maximum (max) value, skewness value and kurtosis value [50]. While it was not necessary to give formulations of energy, median, mean, std, min, max, skewness and kurtosis values that are commonly known from these properties, descriptive information of Shannon entropy, log energy entropy and TKEO characteristics were given below.

Entropy, a measure of random or irregularity of the data being studied, is a nonnegative value [59]. It has a positive value starting with zero and increases in direct proportion to the irregularity in the data [59]. In the theory of knowledge, Shannon entropy (E) is described below with Formula 1 [59]. In Formula 1, P_i represents the probability of the $data_i$ [59].

$$E = -\sum_{i=1}^{N} P_i \log_2 P_i. \tag{1}$$

The equation of the log energy entropy (H) feature is shown in Formula 2 [60].

$$H(x) = -\sum_{i=1}^{N-1} \left(\log_2 \left(P_i(x)\right)\right)^2. \tag{2}$$

Monitor energy in audio signals is used by TKEO method and in which three adjacent data can be calculated in real time [61–65]. A discrete TKEO equation was given by Formula 3. In Formula 3; x is data and n is the sample number.

$$\psi[x(n)] = x^2(n) - x(n+1)\,x(n+1). \tag{3}$$

More detailed and important information on all used features was given in [50].

2.2 Simple Logistic Hybrid System Based on Greedy Stepwise Search Algorithm (SLGS)

An automatic diagnosis system designed on the basis of machine learning techniques requires an elimination mechanism for ineffective attributes and a classification or clustering process to separate the data entered according to the specified classes. These systems can perform highly accurate diagnoses, similar to a specialist, as long as they are adequately trained, equipped with appropriate criteria and selected according to the input [66]. In this study, SLGS system which is a new approach was presented for feature analyzing to diagnose PD according to gender. Primarily, k-fold cross-validation (CV) method was used for each data to divide into tenfold and then created training/testing data were given the proposed SLGS system. Classes were labeled for healthy and patient as 0 and 1, respectively. The accuracy of the performance results obtained from the automatic analysis system was evaluated mathematically and statistically. TP rate (also called the sensitivity, or the recall rate in some fields), FP rate, precision, F, MCC and classification accuracy rate (ACC) [67–72] were used to obtain statistically valid results. Also, ROC and PRC areas [73] were calculated and their curves were plotted additionally.

The design of this system was carried out in a way that feature analysis could be performed according to greedy stepwise search algorithm. It is aimed to achieve maximum success with the number of minimum feature, and Simple Logistic algorithm was chosen for this. The diagram of the SLGS hybrid analysis system is shown in Fig. 2.

When Fig. 2 is examined, feature analysis mechanism of the designed hybrid SLGS system evaluates feature subsets on training data or a separate hold out testing set. This mechanism adopts reduction of irrelevant features like many feature selection methods. The candidate attribute subsets are created by a specific search strategy according to subset generation. [74]. Greedy Stepwise algorithm (with forward direction and none attribute) is used for search strategy in the feature selection mechanism of SLGS hybrid system. This algorithm always selects the next piece by providing the most comprehensible and immediate benefit and thus creates a piecemeal solution. It may start from a random point in space with all attributes or none and searches for feature subsets in the forward or backward direction. The

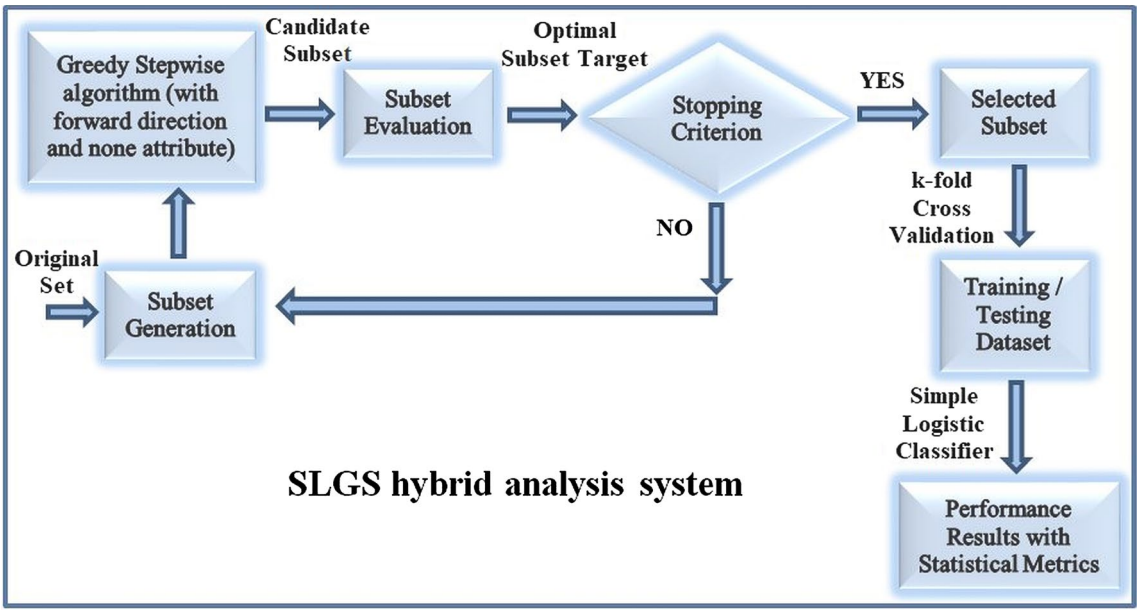

Fig. 2 The diagram of the SLGS hybrid analysis system

algorithm stops the search if there is a decrease in performance results when adding or deleting new features. As a result, it can create the ranked feature list by traversing the space from one side to another [75]. After the search process, a specific evaluation criterion is used to evaluate each candidate subset, and the subsets evaluated are compared to the previous ones based on this criterion. The previous best subset is replaced with a better new subset. The subset generation and evaluation process is repeated until a specific stop criterion is reached. The evaluation measures used by this approach against feature subsets, taking into account the presence and impact of unnecessary features, are defined as being different from individual evaluations. Thus, the optimal subset target is approached with a selected feature subset by this approach [76]. The SLGS hybrid system uses the SL classifier for classification after the effective features are selected. This classifier was created for building linear logistic regression models. Additive logistic regression with simple regression functions is used for fitting the logistic models. In this method, one simple linear regression model is created for each class value [77]. Calculated equation for class probabilities was given by Formula 4 [77].

$$p\left(y_j|x\right) = \frac{e^{F_j(x)}}{\sum_{k=1}^{J} e^{F_k(x)}}, \quad \sum_{k=1}^{J} F_k(x) = 0. \tag{4}$$

In Formula 4, $F_j(x)$ is the linear model for class j, $F_k(x)$ is the models for class j and other classes.

The optimal number of this regression's iterations to perform is cross-validated [78]. The pseudocode of the proposed SLGS hybrid analysis system was given in Fig. 3.

3 Experimental Results

The features extracted from the speech signals of 252 subjects, including 188 PD patients (107 men and 81 women) and 64 healthy (23 men and 41 women) individuals were used for this study [50]. These individuals were made to repeat /a/ letter three times under physician supervision, and the 756 (252×3) data [50] were recorded. The total number of records, separated by male and female genders, were divided into 390 (130×3) and 366 (122×3), respectively. Various features (752 properties in total) were extracted from these recorded data with different techniques by the creators [50] and were categorized into six feature groups used for this study.

SLGS system:
BEGIN
Input: Whole data set
Initialize subset generation
Algorithm 1: Greedy Stepwise algorithm (with forward direction and none attribute)
INPUT: D is initial dataset, F is finishing subset, H is temporary holder, $F_{previous\_result}$
OUTPUT: return optimal subset as F
BEGIN
1. $D = [a_1, a_2, \ldots, a_n]$ where a is attribute; $F = [\,]$
2. H = select a random point in S
3. **FOR** j 1 to n **DO**
   $H \leftarrow F \bigcup$ each possible attribute (candidate subset)
   $C_j \leftarrow$ evaluate each candidate subset H_j
4. determine H_k and C_k with highest evaluation value in H and C
   $F_{new\_result} \leftarrow C_k$
5. **DO** stopping criterion
   IF $F_{new\_result} > F_{previous\_result}$
   $F \leftarrow$ attributes in H_k
   GOTO step 3
   ELSE **RETURN** F

END
GET selected subset $\{f_1, f_2, \ldots, f_n\} \leftarrow F$
Algorithm 2: k-fold cross validation and simple logistic classifier (SL)
BEGIN
1. **DO** randomly partition the selected subset
2. **SET** traindata and testdata
3. SL classifier train model on traindata
4. **GET** model coefficients as $\beta_0, \beta_1, \beta_2, \ldots, \beta_n$
5. **GET** model function as $y = \beta_0 + \beta_1 x f_1 + \beta_2 x f_2 + \ldots + \beta_n x fn$
6. apply model function on testdata
7. obtain test results
8. create a confusion matrix
9. evaluate performance results with statistical metrics

END
END

Fig. 3 The pseudocode of the proposed SLGS hybrid analysis system

At the beginning of the study, the training-test data obtained by a tenfold CV method were given to SLGS hybrid system for the two genders, separately. Then, the SLGS hybrid system started to work and the feature selection process was realized according to the Greedy Stepwise search algorithm for all feature groups. After this process, the selected features with the best results obtained on the basis of gender were given to the SL classifier, which is in the SLGS hybrid system. Finally, the statistical performance metrics of the system were recorded and interpreted. As a result of the system, model function and coefficients obtained for each gender and each class were given by Formula 5 (for males) and Formula 6 (for females) below.

$$
\begin{aligned}
y_{\text{male}} &= \beta_0 + [\text{DFA}] * \beta_1 \\
&\quad + [\text{RPDE}] * \beta_2 + [\text{meanPeriodPulses}] * \beta_3 \\
&\quad + [\text{stdDevPeriodPulses}] * \beta_4 + \left[\text{energy6}\right] * \beta_5 \\
&\quad + \left[\text{energy11}\right] * \beta_6 + \left[\text{energy18}\right] * \beta_7 \\
&\quad + \left[\text{entropyenergylog28}\right] * \beta_8 + [\text{skewnessValue32}] * \beta_9 \\
&\text{for class } 0 \rightarrow \beta_0 = -0.22 \quad \beta_1 = -4.86 \\
&\quad \beta_2 = -3.68 \quad \beta_3 = 383.85 \quad \beta_4 = 101.07 \\
&\quad \beta_5 = 140.64 \quad \beta_6 = 78.14 \quad \beta_7 = 2.81 \\
&\quad \beta_8 = -0 \quad \beta_9 = -0.1 \\
&\text{for class } 1 \rightarrow \beta_0 = 0.22 \quad \beta_1 = 4.86 \quad \beta_2 = 3.68 \\
&\quad \beta_3 = -383.85 \quad \beta_4 = -101.07 \quad \beta_5 = -140.64 \\
&\quad \beta_6 = -78.14 \quad \beta_7 = -2.81 \quad \beta_8 = 0 \quad \beta_9 = 0.1
\end{aligned} \tag{5}
$$

$$
\begin{aligned}
y_{\text{female}} &= \beta_0 + [\text{meanMFCC0}] * \beta_1 + [\text{meanMFCC2}] * \beta_2 \\
&\quad + [\text{meanMFCC11}] * \beta_3 + \left[\text{stddeltalogenergy}\right] * \beta_4 \\
&\quad + \left[\text{entropylog11}\right] * \beta_5 + [\text{TKEOstd18}] * \beta_6 \\
&\quad + [\text{stdValue8}] * \beta_7 + [\text{stdValue22}] * \beta_8 + [\text{kurtosisValue36}] * \beta_9 \\
&\text{for class } 0 \rightarrow \beta_0 = 4.35 \quad \beta_1 = -0.14 \quad \beta_2 = -0.27 \\
&\quad \beta_3 = -0.23 \quad \beta_4 = -27.65 \quad \beta_5 = 0 \\
&\quad \beta_6 = -5.91 \quad \beta_7 = 17.83 \quad \beta_8 = 1.03 \quad \beta_9 = -0.01 \\
&\text{for class } 1 \rightarrow \beta_0 = -4.35 \quad \beta_1 = 0.14 \quad \beta_2 = 0.27 \\
&\quad \beta_3 = 0.23 \quad \beta_4 = 27.65 \beta_5 = -0 \quad \beta_6 = 5.91 \\
&\quad \beta_7 = -17.83 \quad \beta_8 = -1.03 \quad \beta_9 = 0.01.
\end{aligned} \tag{6}
$$

These obtained metrics and the best features (on the basis of gender) are shown in Table 1. The selected final effective features in which the best results were obtained are given in the last column of Table 1.

According to the results in Table 1, 346 inputs of the 390 were correctly labeled and 88.71% ACC ratio was obtained in the diagnosis of PD for male subjects. In addition, while the TP rate and ROC area criteria were closer to value of "1," the FP rate criterion approached the value of "0," and they supported the ACC rate statistically. Besides, the 87.15% ACC ratio was obtained for female subjects and 319 of the total 366 inputs were classified correctly. Similarly, the TP rate criterion approached the value of "1" and they supported the ACC rate statistically. In addition, the FP rate criterion was approached to value of "0" and the success was statistically confirmed. It was seen that much better statistical results were obtained with much less features for TP rate, FP rate, precision, F and MCC criteria. This situation shows the success of the study. Classification errors of SLGS hybrid system for male and female subjects are shown in Fig. 4. In this figure, the blue shapes represented samples of healthy subjects, while the reds indicated samples of patients. When Fig. 4 was examined for both genders, the correct classified data were represented by x, while the wrong ones were indicated by square shapes.

As seen in Fig. 5, ROC and PRC areas criteria for both genders approached the "0.9" band for last effective attributes with SLGS hybrid system. As is known, the area under the ROC curve is the most commonly used ROC statistic. The total area under the ROC curve is the performance measure of diagnostic tests since it reflects test performance at all possible thresholds. In addition to this statistic, precision and recall are inversely related and a balance must be achieved between these metrics. So, precision is in sense correctness, and recall is completeness. The PRC curve shows the balance between these two metrics. As a result, obtained performance metrics of ROC and PRC areas once again proved the success of classification as explained in above the information.

Lastly, when selected features are examined for gender-based diagnosis of PD, intensity parameters and bandwidth features were found the least efficacious feature groups for male and female subjects, respectively. However, it was seen that baseline, MFCCs and the TQWT features had higher likelihood of distinguishing patients and healthy subjects than the other groups for gender analysis. The distribution of the effective attributes determined by gender base using SLGS hybrid system is shown in Figs. 6 and 7.

The numerical range of each selected feature and how the classes were separated from each other are shown with Figs. 6 and 7. In these figures, the area indicated in blue specified the data of the healthy persons and the patients indicated in red. In addition, number of data for each class could be seen through these two figures. According to Table 1, Figs. 6 and 7, it was seen that the number of attributes with the best statistical results was 11 and 9 for male and female subjects, respectively. When these two figures were examined, it was seen that the distinction between patient and healthy classes was generally significant for each feature.

4 Discussion

PD progresses slowly and occurs with a loss of brain cells [1]. The incidence of this disease in male subjects is higher than in female subjects, and one in every 100 people over 65 years of age is considered to have PD [1]. Although there are many symptoms of this disease, approximately 90% of the patients have voice and pronunciation disorders, which are the most important symptoms that are found in the early stages. Previous studies have found that speech recordings are mostly used by the systems in automatic detection of PD [21–24, 26, 27, 40, 50, 79, 80]. This can be diagnosed by means of a specialist physician from the audio signals received from subjects or by automatic diagnostic systems

Table 1 Statistical performance metrics of the SLGS system by gender with determined effective features

Statistical parameters	TNI	CCI	TP rate	FP rate	Precision	F	MCC	ACC (%)	ROC	PRC	Selected best features according to gender
Male	390	346	0.88	0.44	0.88	0.87	0.56	88.71	0.853	0.910	DFA, RPDE, mean period pulses, std dev period pulses, energy (6, 11 and 18th sub-bands), entropy energy log (28th sub-band), mean value (17 and 31st sub-bands), skewness value (32nd sub-band)
Female	366	319	0.87	0.18	0.87	0.87	0.70	87.15	0.884	0.889	Mean MFCC (0, 2 and 11th coefficients), std delta log energy, entropy log (11th sub-band), TKEO std (18th sub-band), std value (8 and 22nd sub-bands), kurtosis value (36th sub-band)

TNI total number of instances, *CCI* correctly classified instances

Fig. 4 Classification errors of SLGS system for male and female subjects

that are created by training artificial intelligence-based algorithms, which have high accuracy and speed. Expert systems can become the tools that people use for a preliminary assessment, especially if they can be integrated into real life in regions that lack medical specialties. Also, by using expert systems in hospitals, a preliminary perspective can be presented to physicians regarding the accuracy of the disease diagnosis through physician decisions. In addition, utilizing the internet, automatic disease diagnostic systems can be used online for individuals, and this situation may play an important role for the early detection of many diseases. In the literature, PD-related engineering studies have generally followed the steps of data recording, data processing, feature extraction from these data, then feature selection and classification. The comparative analysis of the previous studies in which PD diagnosis was made using voice recordings is given in Table 2. This table is generally evaluated in terms of the number of subjects, the methods used and the resulting accuracy rates. In previous works, the obtained accuracy rates vary between 82.5 and 100%. Briefly, the common aim of the studies in the literature is to classify PD with high accuracy rates. While the number of subjects in this paper and study [50] is 252, the number of subjects in other studies varies between 31 and 50.

After reviewing the literature, the greatest shortcoming in such studies was found to be the inability to adequately test the validity of the proposed systems due to limited data. Furthermore, the deficiencies in the recording process adversely affected both the quality of the data and system performance. In these studies, ANN, SVM, KNN, RF classifiers and signal processing techniques such as wavelet, Fourier, principal component analysis (PCA) and linear discriminant analysis (LDA), which are frequently used in the literature, were preferred. In other words, there are few studies—unlike mine—in which a new approach is proposed. Another shortcoming of studies in this field is the limited number of features extracted from the available data due to the lack of signal formats of the data in ready-made data banks, such as UCI. Lastly, the missing aspect of these studies carried out for this purpose is the lack of a detailed analysis of the features

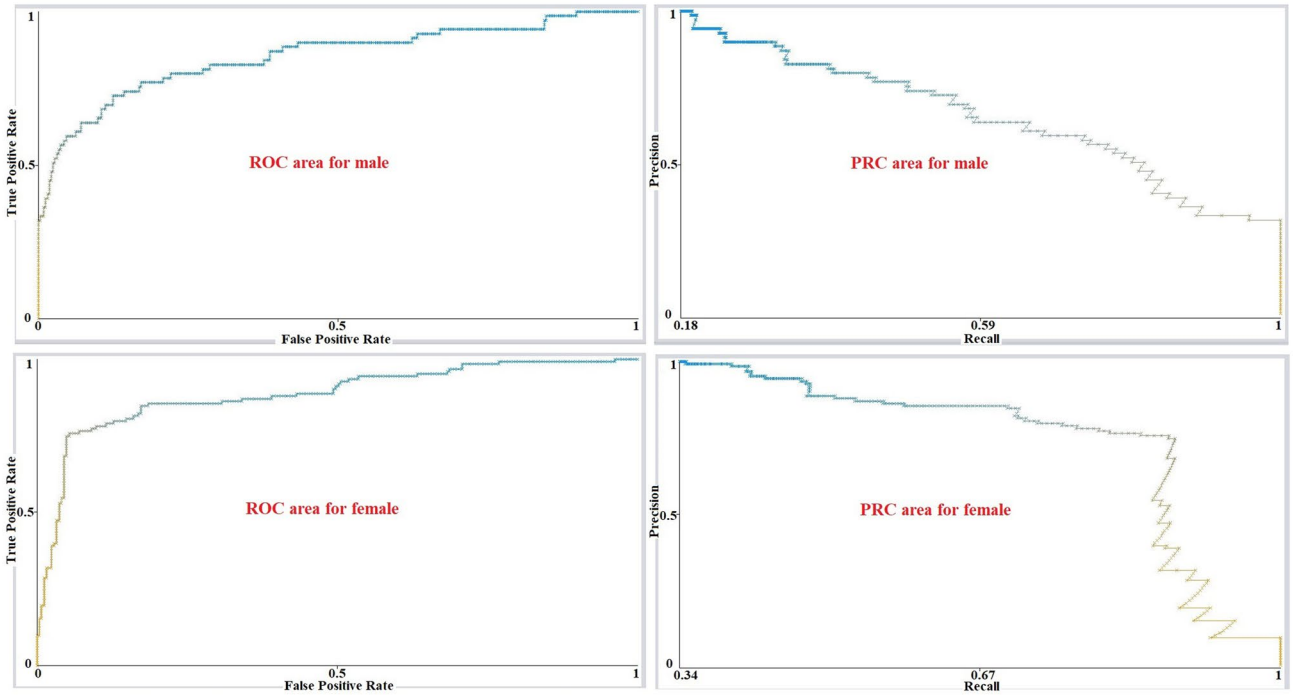

Fig. 5 ROC and PRC areas of the selected best features according to gender with SLGS hybrid system

extracted from the recorded data by gender. Thanks to this study, in order to resolve this deficiency in the literature, the features obtained from the speech recordings were analyzed with high accuracy on the basis of gender. Moreover, in the literature, a detailed analysis about the features extracted from the sound signals according to gender with greedy search algorithm was not completed; and the fact that the number of subject used in this study is quite high (compared to similar studies) once again reveals the importance of this work.

As seen in Table 2, the only previously published study that was used as a data source here is [50]. For this reason, the performance results of the proposed hybrid SLGS system were compared with study [50]. The groups of features mentioned in [50] were classified with different classifiers, and the classifier with the best results for each group of features was generally modified. This situation is a limitation for an expert system as a result of increasing the time and processing intensity. Besides, if the system does not have a specific classifier, it may come to a standstill in terms of operations. In this study, the classifier option was initially reduced to one and the SL classifier was selected. Secondly, in the present study [50], the general performance results of the formed feature groups were investigated, but the analysis of the effective features by gender was not carried out. However, the most effective features of the best results obtained in PD diagnosis on the basis of gender were especially focused on and attempted to reduce the data density of the hybrid SLGS system. Finally, in [50], the minimum redundancy-maximum relevance (mRMR) feature selection method was applied to the whole dataset, and a maximum of 86% ACC ratio was obtained with 50 selected features. In the current study, however, 11 features in males and 9 features in females were selected to generate the effect of the greedy search algorithm and resulted in higher accuracy rates (88.71% for males, 87.15% for females) than the performance results in [50]. When the results were evaluated in terms of other statistical criteria other than ACC, TP rate, FP rate, precision, F, MCC, ROC and PRC values were obtained as 0.88, 0.44, 0.88, 0.87, 0.853, 0.910 for males and 0.87, 0.18, 0.87, 0.87, 0.70, 0.884, 0.889 for females, respectively. This situation is made this study superior to the other one [50] in terms of performance, accuracy, detailed analysis system and the effectiveness of the automatic detection system. In addition, the statistical criteria, such as precision, ROC and PRC areas, were used for the interpretation of the classification results, which supported the success of this study, in contrast to [50]. As a result, the new approach proposed in this paper is expected to make a significant contribution to this field. Even a 0.1% increase in the success of automatic diagnostic systems can likely be of important use in the medical field. Briefly, this proposed method will provide better performance results with a minimum number of features.

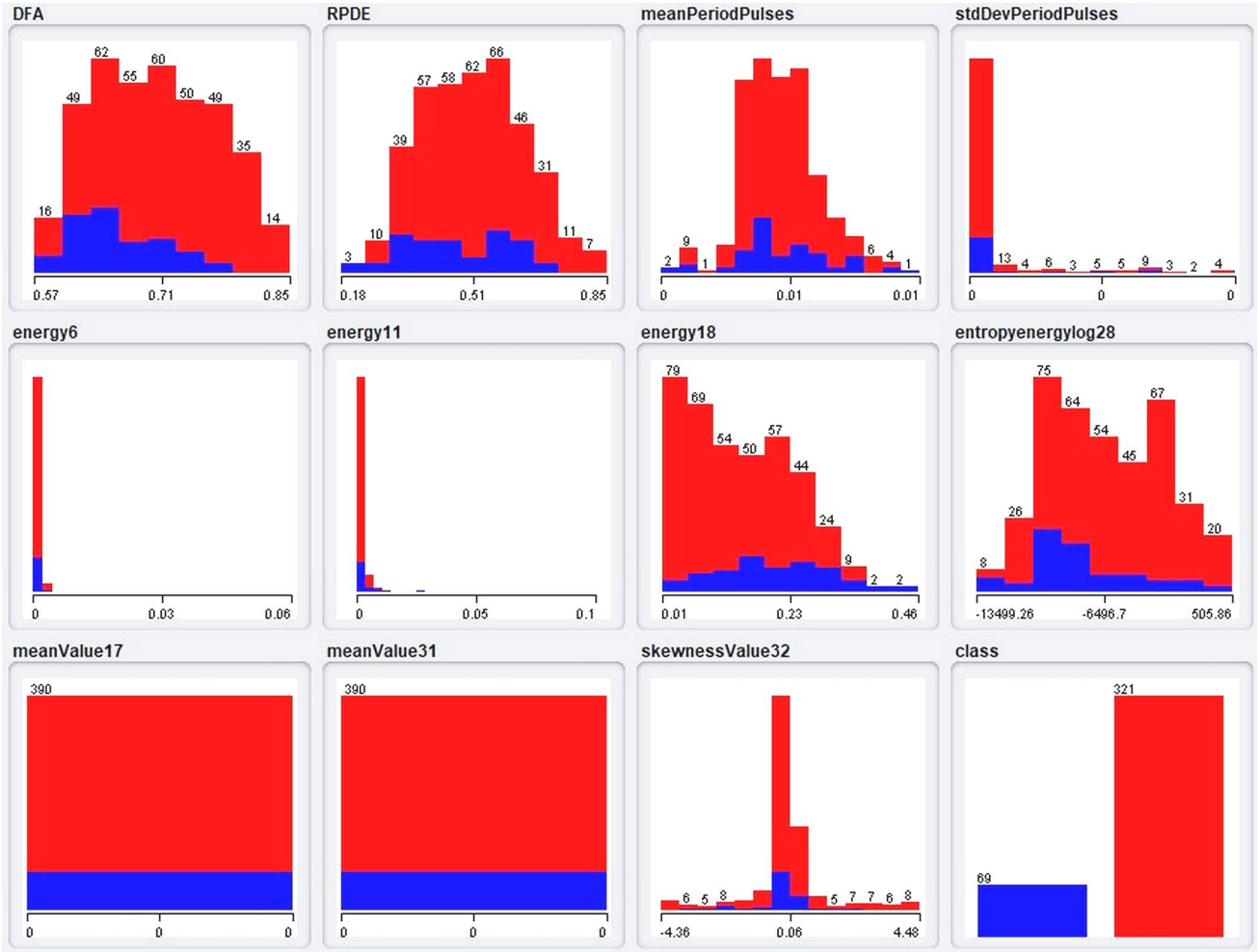

Fig. 6 Distribution of the effective attributes found by SLGS hybrid system for males

The number of attributes that determine the current size of the data affects the computation time of the proposed analysis system required for its implementation. Therefore, while the number of dimensions used in the study was initially 752, this was reduced to 9 and 11 for males and females, respectively, by the greedy stepwise algorithm. When the computation time of the proposed analysis system was evaluated, 2 min 10 s and 3 min 30 s were obtained for females and males, respectively. In the next section, the calculation time of the system on the refined data was recorded as 0.04 s for women and 0.05 s for men. These computation times are quite reasonable with respect to determining and classifying effective properties by analyzing an extremely large dataset. As with every study, the presented study has some limitations. The first is that it used datasets from the UCI database, which consists of extracted features and does not have a signal format. This means that a more detailed analysis of the data could not be performed because there was no access to the signal. Despite this limitation, the performance results obtained as a system output are remarkably successful relative to those of other studies. Furthermore, the use of a single microphone during the recording phase of the audio data used in this study may have prevented a clean capture of the signals, which is believed to have negatively affected the performance results obtained as a result of the system. Moreover, the quality of the voice recorder used in recording operations can affect the data obtained and hence affect the results. Additionally, the proposed hybrid system was not designed in accordance with incremental learning; consequently, it narrowed the extent of the inferences that could be made. When the studies in the literature are examined, most of them try the systems they propose on preexisting datasets. The methods used in the literature are, therefore, often unsuitable for incremental learning. However, if the SLGS system used in this study and the other systems used in the literature are implemented in online environments, the training of the models should be continuously updated by enabling new data entry into the system. This means that the training phases of the systems are made dynamic/adaptive. Future studies should consider that the

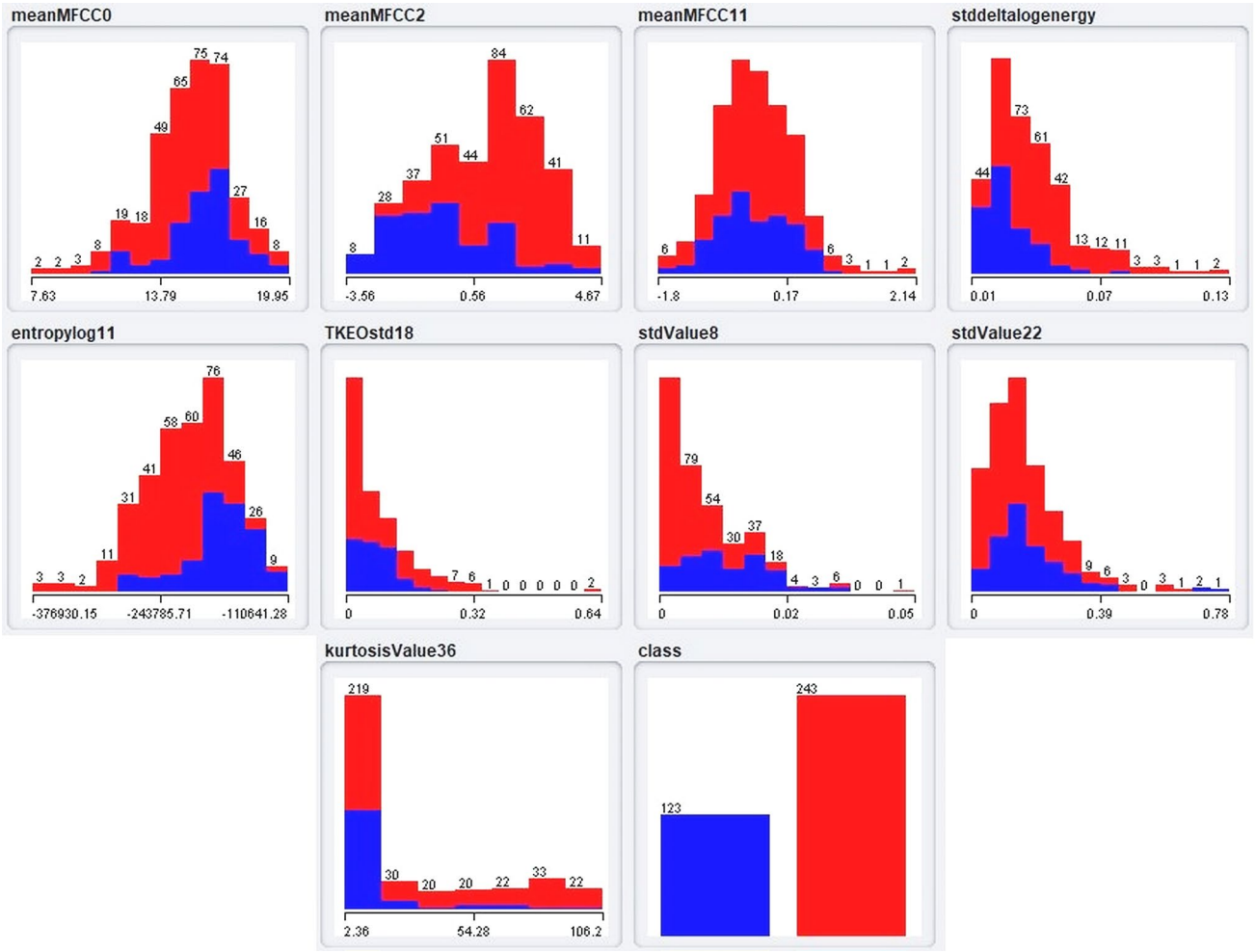

Fig. 7 Distribution of the effective attributes found by SLGS hybrid system for females

SLGS analysis system presented in this study is suitable for adaptation. Finally, the classification performance of the proposed hybrid system must be evaluated with more data in future studies. This limitation will be largely eliminated thanks to the implementation plan of the proposed SLGS system on different medical data of varying sizes.

5 Conclusion

In machine-based self-diagnostic applications, feature extraction, preprocessing, selection and reduction are extremely important and necessary steps after the recorded data are cleaned from unwanted components, such as noise. The common goal shared by all these steps is to improve the performance of the classifier system by reducing system complexity. In this study, a new hybrid system approach is presented, which has not previously been applied to any data. The proposed approach is defined as SLGS; this system performs feature analysis on data and performs automatic PD diagnosis according to gender. The main innovation of this hybrid system is the automatic efficiency analysis of the features extracted from the audio signals for the gender-based diagnosis of PD. The efficiency analysis of features in the system is performed by greedy stepwise search algorithm, which achieves the highest performance results with the fewest number of effective features. In the scope of the study, 11 effective features for men and 9 for women were determined by the greedy search algorithm.

In the next stage of this system, success rates of 88.71% for men and 87.15% for women were achieved. In addition, the results were evaluated according to ROC and PRC criteria, and both metrics approximated 0.9. This result proves that data from PD patients and from healthy subjects are nearly perfectly differentiated from each other. To briefly summarize, a maximum success rate was achieved with a minimum number of features thanks to the generated SLGS hybrid system. Because there are no other studies that create

Table 2 Comparison of the obtained results with the studies in the literature

Study	Number of subjects	Used method	Used classifier system	Accuracy (%)
Sakar and Kursun [21]	32 (PD: 24 Healthy: 8)	Mutual information + mRMR (maximum relevance–minimum redundancy)	Support vector machine (SVM)	92.75
Gürüler [22]	31 (PD: 23 Healthy: 8)	k-Means clustering-based feature weighting (KMCFW)	And a complex-valued artificial neural network (CVANN)	99.52
Little et al. [23]	31 (PD: 23 Healthy: 8)	Pitch period entropy (PPE)	Kernel-SVM	91.4
Peker et al. [24]	31 (PD: 23 Healthy: 8)	mRMR attribute selection algorithm	CVANN	98.12
Sakar et al. [26]	50 (PD: 42 Healthy: 8)	Unified Parkinson's Disease Rating Scale (UPDRS) score method	Kernel-SVM	96.4
Cai et al. [40]	31 (PD: 23 Healthy: 8)	Relief feature selection method	SVM based on bacterial foraging optimization (BFO)—BFOSVM	97.42
Benba et al. [27]	40 (PD: 20 Healthy: 20)	Mel-frequency cepstral coefficients (MFCCs) technique	Kernel-SVM	82.5
Sadek et al. [79]	31 (PD: 23 Healthy: 8)	Feature selection	Artificial neural network system (ANN)	100
Berus et al. [80]	40 (PD: 20 Healthy: 20)	Filter-based feature selection method (A-MCFS)	ANN	86.47
Sakar et al. [50][a]	252 (PD: 188 Healthy:64)	Tunable Q-factor wavelet transform (TQWT) and mRMR-50	SVM (RBF kernels)	86 (with 50 features)
This study[a]	252 (PD: 188 Healthy:64)	A new approach: Simple Logistic hybrid system based on greedy stepwise algorithm (SLGS)		88.71 (for males with 11 features) 87.15 (for females with 9 features)

[a]In these studies, same dataset was used

and test an engineering-based diagnosis method for PD according to gender, this study is novel and important. In addition, the number of subjects was higher in this study than in others.

Further studies using the same proposed hybrid system can be applied to combine the models of HandPD, GaitPD, SpeechPD and Voice datasets. The results of the SLGS system can be further improved with larger PD samples and more meaningful features. In addition, the existing performance rates can be increased by applying feature reduction methods, such as PCA and LDA, on the data properties presented to the system input. Furthermore, the search algorithm and classifier parts within the proposed SLGS system can be modified and run for this dataset or for others.

In the future, the proposed SLGS system is planned to be implemented on other medical—or other kinds of—datasets. Also, the sensitivity analysis of this system can be examined by individually varying, within a certain range, each of the input characteristics determined based on gender to the result. Sample-based methods (input–output correlation or standard regression coefficients) can be used to perform this analysis. In this way, property value ranges can be determined in which the result will not be significantly affected. However, in these and similar automated medical studies, it is sometimes difficult to estimate the value ranges of the properties of the data to be added since the system entries comprise real—rather than hypothetical—data. As a result of sensitivity analysis, value ranges are determined for any input variable, which does not affect the system result. If any real-life variable beyond the specified range is included in the system, system performance may be positively or negatively affected; that is, the performance result may exceed the previously determined range.

Funding No funding was used for this study.

Compliance with Ethical Standards

Conflict of interest The author declares that she has no conflict of interest.

References

1. Jankovic, J.: Parkinson's disease: clinical features and diagnosis. J. Neurol. Neurosurg. Psychiatry **79**(4), 368–376 (2008)

2. Schrag, A.; Anastasiou, Z.; Ambler, G.; Noyce, A.; Walters, K.: Predicting diagnosis of Parkinson's disease: a risk algorithm based on primary care presentations. Mov. Disord. **34**(4), 480–486 (2019)
3. Foulds, P.G.; Mitchell, J.D.; Parker, A.; Turner, R.; Green, G.; Diggle, P.; Hasegawa, M.; Taylor, M.; Mann, D.; Allsop, D.: Phosphorylated α-synuclein can be detected in blood plasma and is potentially a useful biomarker for Parkinson's disease. FASEB J. **25**(12), 4127–4137 (2011)
4. Anita, S.; Priya, P.A.: Diagnosis of Parkinson's disease at an early stage using volume rendering SPECT image slices. Arab. J. Sci. Eng. (2019). https://doi.org/10.1007/s13369-019-04152-7
5. Miller, I.N.; Cronin-Golomb, A.: Gender differences in Parkinson's disease: clinical characteristics and cognition. Mov. Disord. **25**(16), 2695–2703 (2010)
6. Dluzen, D.; McDermott, J.: Gender differences in neurotoxicity of the nigrostriatal dopaminergic system: implications for Parkinson's disease. J. Gend. Specif. Med. **3**(6), 36–42 (2000)
7. Van Den Eeden, S.K.; Tanner, C.M.; Bernstein, A.L.; Fross, R.D.; Leimpeter, A.; Bloch, D.A.; Nelson, L.M.: Incidence of Parkinson's disease: variation by age, gender, and race/ethnicity. Am. J. Epidemiol. **157**(11), 1015–1022 (2003)
8. Haaxma, C.A.; Bloem, B.R.; Borm, G.F.; Oyen, W.J.; Leenders, K.L.; Eshuis, S.; Booij, J.; Dluzen, D.E.; Horstink, M.W.: Gender differences in Parkinson's disease. J. Neurol. Neurosurg. Psychiatry **78**(8), 819–824 (2007)
9. Bordelon, Y.; Fahn, S.: Gender differences in movement disorders. In: Kaplan, P.W. (ed.) Neurologic Disease in Women, pp. 349–354. Demos Medical Publishing, New York (2006)
10. Lavalaye, J.; Booij, J.; Reneman, L.; Habraken, J.B.; van Royen, E.A.: Effect of age and gender on dopamine transporter imaging with [123 I] FP-CIT SPET in healthy volunteers. Eur. J. Nucl. Med. **27**(7), 867–869 (2000)
11. Mozley, L.H.; Gur, R.C.; Mozley, P.D.; Gur, R.E.: Striatal dopamine transporters and cognitive functioning in healthy men and women. Am. J. Psychiatry **158**(9), 1492–1499 (2001)
12. Munro, C.A.; McCaul, M.E.; Wong, D.F.; Oswald, L.M.; Zhou, Y.; Brasic, J.; Kuwabara, H.; Kumar, A.; Alexander, M.; Ye, W.: Sex differences in striatal dopamine release in healthy adults. Biol. Psychiat. **59**(10), 966–974 (2006)
13. Scott, B.; Borgman, A.; Engler, H.; Johnels, B.; Aquilonius, S.: Gender differences in Parkinson's disease symptom profile. Acta Neurol. Scand. **102**(1), 37–43 (2000)
14. Hariz, G.M.; Lindberg, M.; Hariz, M.I.; Tommy Bergenheim, A.: Gender differences in disability and health-related quality of life in patients with Parkinson's disease treated with stereotactic surgery. Acta Neurol. Scand. **108**(1), 28–37 (2003)
15. Accolla, E.; Caputo, E.; Cogiamanian, F.; Tamma, F.; Mrakic-Sposta, S.; Marceglia, S.; Egidi, M.; Rampini, P.; Locatelli, M.; Priori, A.: Gender differences in patients with Parkinson's disease treated with subthalamic deep brain stimulation. Mov. Disord. **22**(8), 1150–1156 (2007)
16. Baba, Y.; Putzke, J.D.; Whaley, N.R.; Wszolek, Z.K.; Uitti, R.J.: Gender and the Parkinson's disease phenotype. J. Neurol. **252**(10), 1201–1205 (2005)
17. Zappia, M.; Crescibene, L.; Arabia, G.; Nicoletti, G.; Bagalà, A.; Bastone, L.; Caracciolo, M.; Bonavita, S.; Di Costanzo, A.; Scornaienchi, M.: Body weight influences pharmacokinetics of levodopa in Parkinson's disease. Clin. Neuropharmacol. **25**(2), 79–82 (2002)
18. Yoritaka, A.; Ohizumi, H.; Tanaka, S.; Hattori, N.: Parkinson's disease with and without REM sleep behaviour disorder: Are there any clinical differences? Eur. Neurol. **61**(3), 164–170 (2009)
19. Harel, B.; Cannizzaro, M.; Snyder, P.J.: Variability in fundamental frequency during speech in prodromal and incipient Parkinson's disease: a longitudinal case study. Brain Cogn. **56**(1), 24–29 (2004)
20. Tsanas, A.; Little, M.A.; McSharry, P.E.; Ramig, L.O.: Accurate telemonitoring of Parkinson's disease progression by noninvasive speech tests. IEEE Trans. Biomed. Eng. **57**(4), 884–893 (2010)
21. Sakar, C.O.; Kursun, O.: Telediagnosis of Parkinson's disease using measurements of dysphonia. J. Med. Syst. **34**(4), 591–599 (2010)
22. Gürüler, H.: A novel diagnosis system for Parkinson's disease using complex-valued artificial neural network with k-means clustering feature weighting method. Neural Comput. Appl. **28**(7), 1657–1666 (2017)
23. Little, M.A.; McSharry, P.E.; Hunter, E.J.; Spielman, J.; Ramig, L.O.: Suitability of dysphonia measurements for telemonitoring of Parkinson's disease. IEEE Trans. Biomed. Eng. **56**(4), 1015–1022 (2009)
24. Peker, M.; Sen, B.; Delen, D.: Computer-aided diagnosis of Parkinson's disease using complex-valued neural networks and mRMR feature selection algorithm. J. Healthc. Eng. **6**(3), 281–302 (2015)
25. Cantürk, İ.; Karabiber, F.: A machine learning system for the diagnosis of Parkinson's disease from speech signals and its application to multiple speech signal types. Arab. J. Sci. Eng. **41**(12), 5049–5059 (2016)
26. Sakar, B.E.; Serbes, G.; Sakar, C.O.: Analyzing the effectiveness of vocal features in early telediagnosis of Parkinson's disease. PLoS ONE **12**(8), e0182428 (2017)
27. Benba, A.; Jilbab, A.; Hammouch, A.: Analysis of multiple types of voice recordings in cepstral domain using MFCC for discriminating between patients with Parkinson's disease and healthy people. Int. J. Speech Technol. **19**(3), 449–456 (2016)
28. Sharma, P.; Sundaram, S.; Sharma, M.; Sharma, A.; Gupta, D.: Diagnosis of Parkinson's disease using modified grey wolf optimization. Cogn. Syst. Res. **54**, 100–115 (2019)
29. Almeida, J.S.; Rebouças Filho, P.P.; Carneiro, T.; Wei, W.; Damaševičius, R.; Maskeliūnas, R.; de Albuquerque, V.H.C.: Detecting Parkinson's disease with sustained phonation and speech signals using machine learning techniques. Pattern Recogn. Lett. **125**, 55–62 (2019)
30. Zhang, Y.: Can a smartphone diagnose parkinson disease? A deep neural network method and telediagnosis system implementation. Parkinson's Dis. **2017**, 1–11 (2017). https://doi.org/10.1155/2017/6209703
31. Zhang, H.-H.; Yang, L.; Liu, Y.; Wang, P.; Yin, J.; Li, Y.; Qiu, M.; Zhu, X.; Yan, F.: Classification of Parkinson's disease utilizing multi-edit nearest-neighbor and ensemble learning algorithms with speech samples. Biomed. Eng. Online **15**(1), 122 (2016)
32. Upadhya, S.S.; Cheeran, A.; Nirmal, J.H.: Thomson multitaper MFCC and PLP voice features for early detection of Parkinson disease. Biomed. Signal Process. Control **46**, 293–301 (2018)
33. Upadhya, S.S.; Cheeran, A.: Performance comparison of regression techniques in predicting parkinson disease severity score using speech features. Biomed. Eng. Appl. Basis Commun. **30**(04), 1850025 (2018)
34. Parisi, L.; RaviChandran, N.; Manaog, M.L.: Feature-driven machine learning to improve early diagnosis of Parkinson's disease. Expert Syst. Appl. **110**, 182–190 (2018)
35. Oung, Q.W.; Muthusamy, H.; Basah, S.N.; Lee, H.; Vijean, V.: Empirical wavelet transform based features for classification of Parkinson's disease severity. J. Med. Syst. **42**(2), 29 (2018)
36. Al Mamun, K.A.; Alhussein, M.; Sailunaz, K.; Islam, M.S.: Cloud based framework for Parkinson's disease diagnosis and monitoring system for remote healthcare applications. Fut. Gener. Comput. Syst. **66**, 36–47 (2017)

37. Hariharan, M.; Polat, K.; Sindhu, R.: A new hybrid intelligent system for accurate detection of Parkinson's disease. Comput. Methods Programs Biomed. **113**(3), 904–913 (2014)
38. Haq, A.U.; Li, J.P.; Memon, M.H.; Malik, A.; Ahmad, T.; Ali, A.; Nazir, S.; Ahad, I.; Shahid, M.: Feature selection based on L1-norm support vector machine and effective recognition system for Parkinson's disease using voice recordings. IEEE Access **7**, 37718–37734 (2019)
39. Chen, H.-L.; Wang, G.; Ma, C.; Cai, Z.-N.; Liu, W.-B.; Wang, S.-J.: An efficient hybrid kernel extreme learning machine approach for early diagnosis of Parkinson′ s disease. Neurocomputing **184**, 131–144 (2016)
40. Cai, Z.; Gu, J.; Chen, H.-L.: A new hybrid intelligent framework for predicting Parkinson's disease. IEEE Access **5**, 17188–17200 (2017)
41. Ali, L.; Zhu, C.; Zhou, M.; Liu, Y.: Early diagnosis of Parkinson's disease from multiple voice recordings by simultaneous sample and feature selection. Expert Syst. Appl. **137**, 22–28 (2019)
42. Abiyev, R.H.; Abizade, S.: Diagnosing Parkinson's diseases using fuzzy neural system. Comput. Math. Methods Med. **2016**, 1–9 (2016). https://doi.org/10.1155/2016/1267919
43. Wu, Y.; Chen, P.; Yao, Y.; Ye, X.; Xiao, Y.; Liao, L.; Wu, M.; Chen, J.: Dysphonic voice pattern analysis of patients in Parkinson's disease using minimum interclass probability risk feature selection and bagging ensemble learning methods. Comput. Math. Methods Med. **2017**, 1–11 (2017). https://doi.org/10.1155/2017/4201984
44. Wang, Y.; Wang, A.-N.; Ai, Q.; Sun, H.-J.: An adaptive kernel-based weighted extreme learning machine approach for effective detection of Parkinson's disease. Biomed. Signal Process. Control **38**, 400–410 (2017)
45. Mostafa, S.A.; Mustapha, A.; Mohammed, M.A.; Hamed, R.I.; Arunkumar, N.; Ghani, M.K.A.; Jaber, M.M.; Khaleefah, S.H.: Examining multiple feature evaluation and classification methods for improving the diagnosis of Parkinson's disease. Cogn. Syst. Res. **54**, 90–99 (2019)
46. Lahmiri, S.; Shmuel, A.: Detection of Parkinson's disease based on voice patterns ranking and optimized support vector machine. Biomed. Signal Process. Control **49**, 427–433 (2019)
47. Lahmiri, S.; Dawson, D.A.; Shmuel, A.: Performance of machine learning methods in diagnosing Parkinson's disease based on dysphonia measures. Biomed. Eng. Lett. **8**(1), 29–39 (2018)
48. Cai, Z.; Gu, J.; Wen, C.; Zhao, D.; Huang, C.; Huang, H.; Tong, C.; Li, J.; Chen, H.: An intelligent Parkinson's disease diagnostic system based on a chaotic bacterial foraging optimization enhanced fuzzy KNN approach. Comput. Math. Methods Med. **2018**, 1–24 (2018). https://doi.org/10.1155/2018/2396952
49. Gupta, D.; Julka, A.; Jain, S.; Aggarwal, T.; Khanna, A.; Arunkumar, N.; de Albuquerque, V.H.C.: Optimized cuttlefish algorithm for diagnosis of Parkinson's disease. Cogn. Syst. Res. **52**, 36–48 (2018)
50. Sakar, C.O.; Serbes, G.; Gunduz, A.; Tunc, H.C.; Nizam, H.; Sakar, B.E.; Tutuncu, M.; Aydin, T.; Isenkul, M.E.; Apaydin, H.: A comparative analysis of speech signal processing algorithms for Parkinson's disease classification and the use of the tunable Q-factor wavelet transform. Appl. Soft Comput. **74**, 255–263 (2019)
51. Peker, M.: A decision support system to improve medical diagnosis using a combination of k-medoids clustering based attribute weighting and SVM. J. Med. Syst. **40**(5), 116 (2016)
52. Tsanas, A.; Little, M.A.; McSharry, P.E.; Spielman, J.; Ramig, L.O.: Novel speech signal processing algorithms for high-accuracy classification of Parkinson's disease. IEEE Trans. Biomed. Eng. **59**(5), 1264–1271 (2012)
53. Murty, K.S.R.; Yegnanarayana, B.: Combining evidence from residual phase and MFCC features for speaker recognition. IEEE Signal Process. Lett. **13**(1), 52–55 (2006)
54. Godino-Llorente, J.I.; Gomez-Vilda, P.; Blanco-Velasco, M.: Dimensionality reduction of a pathological voice quality assessment system based on Gaussian mixture models and short-term cepstral parameters. IEEE Trans. Biomed. Eng. **53**(10), 1943–1953 (2006)
55. Kapoor, T.; Sharma, R.: Parkinson's disease diagnosis using Mel-frequency cepstral coefficients and vector quantization. Int. J. Comput. Appl. **14**(3), 43–46 (2011)
56. Frail, R.; Godino-Llorente, J.; Saenz-Lechon, N.; Osma-Ruiz, V.; Fredouille, C.: MFCC-based remote pathology detection on speech transmitted through the telephone channel. In: BIOSIGNALS 2009 - International Conference on Bio-inspired Systems and Signal Processing, pp. 41–48 (2009)
57. Murphy, P.J.; Akande, O.O.: Quantification of glottal and voiced speech harmonics-to-noise ratios using Cepstral-based estimation. In: ISCA Tutorial and Research Workshop (ITRW) on Non-linear Speech Processing (2005)
58. Selesnick, I.W.: Wavelet transform with tunable Q-factor. IEEE Trans. Signal Process. **59**(8), 3560–3575 (2011)
59. Gray, R.M.: Entropy and Information. In Entropy and Information Theory. Springer, New York (1990)
60. Aydın, S.; Saraoğlu, H.M.; Kara, S.: Log energy entropy-based EEG classification with multilayer neural networks in seizure. Ann. Biomed. Eng. **37**(12), 2626 (2009)
61. Kaiser, J.F.: On a simple algorithm to calculate the 'energy' of a signal. In: 1990 International Conference on Acoustics, Speech, and Signal Processing, 1990 (ICASSP-90), pp. 381–384. IEEE (1990)
62. Kaiser, J.F.: Some useful properties of Teager's energy operators. In: IEEE International Conference on Acoustics, Speech, and Signal Processing, 1993 (ICASSP-93). IEEE, pp. 149–152 (1993)
63. Maragos, P.; Kaiser, J.F.; Quatieri, T.F.: On amplitude and frequency demodulation using energy operators. IEEE Trans. Signal Process. **41**(4), 1532–1550 (1993)
64. Solnik, S.; Rider, P.; Steinweg, K.; DeVita, P.; Hortobágyi, T.: Teager–Kaiser energy operator signal conditioning improves EMG onset detection. Eur. J. Appl. Physiol. **110**(3), 489–498 (2010)
65. Randall, R.B.; Smith, W.A.: Application of the Teager–Kaiser energy operator to machine diagnostics. In: Tenth Dst Group International Conference on Health and Usage Monitoring Systems (2017)
66. Santra, A.; Christy, C.J.: Genetic algorithm and confusion matrix for document clustering. Int. J. Comput. Sci. Issues **9**(1), 322 (2012)
67. Ma, Y.; Guo, L.; Cukic, B.: A statistical framework for the prediction of fault-proneness. In: Advances in Machine Learning Applications in Software Engineering. IGI Global, pp. 237–263 (2007)
68. Yücelbaş, Ş.; Yücelbaş, C.; Tezel, G.; Özşen, S.; Yosunkaya, Ş.: Automatic sleep staging based on SVD, VMD, HHT and morphological features of single-lead ECG signal. Expert Syst. Appl. **102**, 193–206 (2018)
69. Yücelbaş, Ş.; Yücelbaş, C.; Tezel, G.; Özşen, S.; Küçüktürk, S.; Yosunkaya, Ş.: Pre-determination of OSA degree using morphological features of the ECG signal. Expert Syst. Appl. **81**, 79–87 (2017)
70. Yücelbaş, C.; Yücelbaş, Ş.; Özşen, S.; Tezel, G.; Küçüktürk, S.; Yosunkaya, Ş.: Automatic detection of sleep spindles with the use of STFT, EMD and DWT methods. Neural Comput. Appl. **29**(8), 17–33 (2018)
71. Dursun, M.; Özşen, S.; Yücelbaş, C.; Yücelbaş, Ş.; Tezel, G.; Küçüktürk, S.; Yosunkaya, Ş.: A new approach to eliminating EOG artifacts from the sleep EEG signals for the automatic sleep

stage classification. Neural Comput. Appl. **28**(10), 3095–3112 (2017)
72. Yücelbaş, C.; Yücelbaş, Ş.; Özşen, S.; Tezel, G.; Küççüktürk, S.; Yosunkaya, Ş.: A novel system for automatic detection of K-complexes in sleep EEG. Neural Comput. Appl. **29**(8), 137–157 (2018)
73. Saito, T.; Rehmsmeier, M.: The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. PLoS ONE **10**(3), e0118432 (2015)
74. Liu, H.; Motoda, H.: Feature Extraction, Construction and Selection: A Data Mining Perspective, vol. 453. Springer, Berlin (1998)
75. Cormen, T.H.; Leiserson, C.E.; Rivest, R.L.; Stein, C.: Introduction to Algorithms. MIT Press, Cambridge (2009)
76. Yu, L.; Liu, H.: Efficient feature selection via analysis of relevance and redundancy. J. Mach. Learn. Res. **5**, 1205–1224 (2004)
77. Friedman, J.; Hastie, T.; Tibshirani, R.: Additive logistic regression: a statistical view of boosting (with discussion and a rejoinder by the authors). Ann. Stat. **28**(2), 337–407 (2000)
78. Landwehr, N.; Hall, M.; Frank, E.: Logistic model trees. Mach. Learn. **59**(1–2), 161–205 (2005)
79. Sadek, R.M.; Mohammed, S.A.; Abunbehan, A.R.K.; Ghattas, A.K.H.A.; Badawi, M.R.; Mortaja, M.N.; Abu-Nasser, B.S.; Abu-Naser, S.S.: Parkinson's disease prediction using artificial neural network. Int. J. Acad. Health Med. Res. **3**(1), 1–8 (2019)
80. Berus, L.; Klancnik, S.; Brezocnik, M.; Ficko, M.: Classifying Parkinson's disease based on acoustic measures using artificial neural networks. Sensors **19**(1), 16 (2019)