International Journal of Speech Technology (2022) 25:583–593
https://doi.org/10.1007/s10772-021-09837-9

A hybrid system for Parkinson's disease diagnosis using machine learning techniques

Rohit Lamba[1] · Tarun Gulati[1] · Hadeel Fahad Alharbi[2] · Anurag Jain[3]

Received: 3 November 2020 / Accepted: 31 March 2021 / Published online: 14 April 2021

Abstract

Parkinson's disease is a neurodegenerative disorder that progresses slowly and its symptoms appear over time, so its early diagnosis is not easy. A neurologist can diagnose Parkinson's by reviewing the patient's medical history and repeated scans. Besides, body movement analysts can diagnose Parkinson's by analyzing body movement. Recent research work has shown that changes in speech can be used as a measurable indicator for early Parkinson's detection. In this work, the authors propose a speech signal-based hybrid Parkinson's disease diagnosis system for its early diagnosis. To achieve this, the authors have tested several combinations of feature selection approaches and classification algorithms and designed the model with the best combination. To formulate various combinations, three feature selection methods such as mutual information gain, extra tree, and genetic algorithm and three classifiers namely naive bayes, k-nearest-neighbors, and random forest have been used. To analyze the performance of different combinations, the speech dataset available at the UCI (University of California, Irvine) machine learning repository has been used. As the dataset is highly imbalanced so the class balancing problem is overcome by the synthetic minority oversampling technique (SMOTE). The combination of genetic algorithm and random forest classifier has shown the best performance with 95.58% accuracy. Moreover, this result is also better than the recent work found in the literature.

Keywords Parkinson's disease · Decision support system · Feature selection · Classifier algorithm · Machine learning

1 Introduction

Parkinson's disease is the second most slowly progressive neurodegenerative disorder after Alzheimer's disease (Sakar et al., 2019). Around seven to ten million people worldwide are affected by this disease (Tuncer et al., 2020). The main cause of Parkinson's is the lack of dopamine in the substantia nigra, which is part of the brain. Dopamine is a brain chemical produced by neurons that work as a neurotransmitter. When neurons in the substantia nigra begin to damage or dies, results in less production of dopamine which is the main cause of the start of Parkinson's in a person. The reason behind the impairment of these neurons is still unknown. Parkinson's disease cannot be cured but an early diagnosis can help the person in proper treatment and avoid the critical situation (Emamzadeh & Surguchov, 2018; Lamba et al., 2020). Levodopa, carbidopa, and syndopa are the commonly used drugs used to control Parkinson's symptoms. These drugs stimulate the left dopamine-producing cells to produce more dopamine. Parkinson's disease can affect both women and men. However, the ratio of Parkinson's disease patients

✉ Rohit Lamba
rohitlamba14@mmumullana.org

Tarun Gulati
gulati_tarun@mmumullana.org

Hadeel Fahad Alharbi
h.alharbe@uoh.edu.sa

Anurag Jain
anurag.jain@ddn.upes.ac.in

[1] Department of Electronics & Communication Engineering, Maharishi Markandeshwar Engineering College, Maharishi Markandeshwar (Deemed To Be University), Mullana, Ambala, Haryana, India

[2] College of Computer Science and Engineering, University of Ha'il, Ha'il, Kingdom of Saudi Arabia

[3] Virtualization Department, School of Computer Science, University of Petroleum & Energy Studies, Dehradun, India

between men and women is 3:2. Parkinson's disease usually develops in people around sixty, but it can develop before the age of fifty (Reich & Savitt, 2018).

The earlier symptoms of Parkinson's disease are very mild and sometimes unnoticeable, but the symptoms become severe as the disease progresses. The symptoms of Parkinson's disease vary from patient to patient. The initial symptoms of Parkinson's are motor and non-motor symptoms. Motor symptoms include bradykinesia, tremor, rigidity, and postural instability (loss of balance) and nonmotor symptoms include psychiatric symptoms, sleep sickness, sensory impairment, and dysautonomia (autonomic dysfunction) (Lamba et al., 2021; Zesiewicz et al., 2019). Change in speech, tremor, slowed movement (bradykinesia), change in handwriting, rigid muscles, impaired posture and balance, loss of automatic movements are the common symptoms of Parkinson's patient. Researchers have found vocal disorder problems in around 90% of Parkinson's patients which is shown at an early stage. These Vocal disorders are hypophonia (reduced volume), dysphonia (defective voice), dysarthria (difficulty with articulation), and monotone (reduced pitch) (Zesiewicz et al., 2019).

The early diagnosis of Parkinson's is not easy due to several reasons. Mainly the neurologists and movement order specialists can diagnose this disease only after reviewing the patient's complete medical history and repeated scans which are both time-consuming and inconvenient for the patients as most of them are above the age of sixty. The domain knowledge of the physicians who analyze the patient's data and symptoms plays a key role in accurate Parkinson's diagnosis. But unfortunately, in developing countries like Argentina, Brazil, India, etc., we do not have a sufficient number of expert doctors. Therefore, the diagnosis or detection of Parkinson's is a challenging task, because experts are under stress due to high workload. This has motivated us to develop a decision support system that can assist Physicians in Parkinson's diagnosis process. This can serve as a second opinion for Parkinson's diagnosis and also reduce the likelihood of errors due to the involvement of machine learning (Jain et al., 2019; Zesiewicz et al., 2019).

Researchers have used several ways such as EEG signals, Freezing of Gait (FoG), EMG signals, SPECT images, MRI images, and handwritten images for Parkinson's diagnosis. Changes in speech called dysphonia, have also been shown as an early symptom that appears in Parkinson's patients, which can be used to diagnosing Parkinson's. Therefore speech changes have been chosen in this study for diagnosis of Parkinson's because it is simple, non-invasive, and low cost. Researchers have implemented several machine learning techniques to build decision support systems (DSS) including pre-processing, feature engineering, classification, and validation steps. Disease patterns can be easily analyzed from medical datasets and decisions can be made much faster using machine learning techniques. Preprocessing techniques include data normalization and balancing. Feature extraction and selection is a process of feature engineering. Feature selection reduces the computational cost and increases accuracy.

The major contribution of this work is as follows:

1. Firstly the feature selection was done by three feature selection methods namely mutual information gain, extra tree, and genetic algorithm and out of 23 only 5 and 11 features were selected. The reduced feature subset was used for training and testing of classifiers to identify the best combination of feature selection method and classifier.
2. Secondly, the Parkinson's disease speech dataset used in this study is highly imbalanced because 147 samples out of 197 samples are from Parkinson's patients, so SMOTE has been used to handle the class imbalance problem.
3. Finally, the performance of three classifiers namely naive bayes, k-nearest neighbors, and random forest was analyzed on full features and reduced feature subsets. The combination of genetic algorithm and random forest classifier has performed better than the decision support systems found in the literature with an accuracy of 95.58%.

The paper is organized as follows. In Sect. 2, an overview of the literature related to Parkinson's disease diagnosis is given. The proposed Methodology and dataset used are given in Sect. 3. The results and their discussion are given in Sect. 4. Conclusion and future scope are outlined in Sect. 5.

2 Literature survey

Researchers have proposed various Parkinson's disease diagnosis systems using speech signals. Naranjo et al., (2016) presented a clinical expert system to detect Parkinson's disease. The voice of eighty individuals including half Parkinson's patients was recorded three times with phonation of /a/ vowels for at least five seconds. Forty-four features from five different categories were extracted by different algorithms such as a waveform matching algorithm and a dataset of 240 rows and 44 columns were prepared. A novel subject-based Bayesian classification approach was used because the dataset has three sound recordings of each individual. After applying the cross-validation, 75.2% accuracy was achieved. Gupta et al. (2018) have used decision tree and KNN classifiers to diagnose Parkinson's using speech, voice, and HandPD datasets. An optimized cuttlefish algorithm was used for feature selection to enhance accuracy by using fewer features. Maximum 92.19% accuracy was obtained from the proposed system.

Sakar et al. (2019) proposed a decision support system for the classification of patients with Parkinson's disease. Using various speech processing algorithms, useful clinical information for the classification of PD patients is extracted and fed to machine learning classifiers. Tunable Q-factor wavelet transform (TQWT) is used the first time for feature extraction from voice signals. The feature selection was done by minimum redundancy-maximum relevance (mRMR) method and classification was done by six different classifiers. The highest 86% accuracy was achieved by the combination of mRMR and SVM-RBF algorithms. Grover et al. (2018) proposed a methodology for Parkinson's disease severity prediction using deep neural network (DNN) and considering the unified Parkinson's disease rating scale (UPDRS score). Two experiments were done for Motor-UPDRS and Total-UPDRS scores. Results showed that the classification accuracy based on the Motor-UPDRS score was 81.66% which was better than the accuracy achieved by the Total-UPDRS score. Lahmiri and Shmuel (2019) have used voice disorder patterns for the diagnosis of Parkinson's patients. Eight different pattern ranking techniques for feature selection and a nonlinear SVM classifier were used to distinguish between healthy and Parkinson's disease patients. An accuracy of 92.21% was achieved with the Wilcox statistic feature rank method.

Almeida et al. (2019) proposed a PD detection system using potation and speech signals. By using Smartphone (SP) and acoustic cardioids (AC) audio signals were recorded. The recorded signals were pre-processed and divided into voiced and unvoiced files by using special software. A total of 144 features were extracted by applying feature engineering techniques. The training of KNN, MLP, and SVM classifiers were done by OpenCV 2.49 Toolbox. The results prove that the potation task was more efficient than the speech task and the accuracy obtained by using the AC channel and SP Channel was 94.55% and 92.94% respectively. Sharma, Jain, et al. (2019) proposed an antlion optimization (ALO) algorithm-based prediction model for Parkinson's disease. The reduced feature subset selected by the proposed ALO algorithm was feed into the KNN, decision tree, and random forest classifiers, and maximum 95.91% accuracy was achieved. Sharma, Sundaram, et al. (2019) proposed a Parkinson's disease diagnosis system in which Parkinson's speech, Parkinson's voice, and HandPD datasets from the UCI machine learning repository were used. The authors used a modified form of the popular grey wolf optimization algorithm for feature selection. Three classifiers KNN, random forest, and decision tree classifiers were used for classification. The highest accuracy of 93.87% was achieved on the speech dataset.

Tuncer and Dogan (2019) proposed a novel octopus-based method for gender reorganization and detection of Parkinson's disease. Features were first extracted by using the singular value decomposition (SVD) method and then feature selection was done by neighborhood component analysis (NCA). Six different classifiers were used for classification. The proposed methodology achieved an accuracy of 99.21% for gender, 98.41% for PD, and 97.62% for PD + Gender. Nissar et al. (2019) proposed a PD detection system by using voice signals. The feature selection was performed by mRMR and RFE methods. The classification was performed by eight classifiers. The results show that the combination of RFE with XGboost outperformed with 95.39% accuracy. Gunduz (2019) proposed two CNN-based frameworks to classify Parkinson's disease. The speech data from the UCI repository was used in this paper. The first framework was named feature-level combination and the second model level combination. The model level framework shows the highest of 86.9% accuracy.

Polat (2019) proposed a hybrid model for Parkinson's diagnosis using speech signals. The dataset was first pre-processed by the SMOTE technique. The classification was performed by random forest classifier and 94.8% accuracy was achieved. Mostafa et al. (2019) proposed a Parkinson disease diagnosis method by analyzing the voice disorders of Parkinson patients. A novel multiple feature evaluation method was used for feature selection. The performance of the reduced feature subset is analyzed by five different classifiers and the random forest classifier outperforms by achieving 99.49% accuracy. Tuncer et al. (2020) presented how vowels can be used for Parkinson's detection. After pre-processing the data using the MAMa tree, feature selection was done by singular value decomposition and relief-based method. The classification was performed by eight classifiers. An accuracy of 92.46% was obtained by using the KNN classifier.

Zahid et al. (2020) proposed a spectrogram-based deep feature extraction method for Parkinson's disease diagnosis. Spanish language PC-GITA data was used in this study. Three different methods were proposed in this study. In the first approach, speech signals were converted into a spectrogram then pre-trained CNN with ALEXNET was used for feature extraction. In the second approach, the same pre-trained CNN model was used for the extraction of features from speech signals. In the third approach simple acoustic signals, spectral and statistical features were extracted and fed into the classifiers. The results show that the average accuracy of 99.3% was obtained by a multilayer perceptron classifier.

Olivares et al. (2020) proposed a Parkinson's diagnosis system by using an optimized version of the BAT algorithm. Only 23 features were selected from UCI Parkinson's disease classification data set and directly feed into the 23 neurons in the input layer of the model. The 96.74% accuracy was achieved by the proposed method with a 3.27% loss. Solana-Lavalle et al. (2020) proposed vocal features based early Parkinson's pre-diagnosis tool. The wrapper method is used for feature selection and RF, MLP, SVM, and KNN classifiers were used for classification. The highest 94.7% accuracy was obtained by using SVM-RBF Classifier.

Yaman et al. (2020) presented a Parkinson's detection model by using vowels. Acoustic features from the dataset were selected by the ReliefF method and classification was done by KNN and SVM classifiers. The highest 91.25% accuracy was achieved by the SVM classifier. Xiong et al. (2020) proposed an adaptive grey wolf optimization algorithm based Parkinson's classification model by using speech signals. Sparse autoencoder was deployed for dimensionality reduction and classification was done by eight classifiers. The performance of states of the art features selection techniques like minimum Redundancy Maximum Relevance (mRMR), correlation-based feature selection (CFS), and recursive feature elimination (RFE) was tested with sparse autoencoders. Senturk (2020) presented a machine learning algorithms-based diagnosis system for Parkinson's disease. Feature selection was done by RFE and feature importance methods. Regression tree, ANN, and SVM were used as classifiers. The combination of RFE with the SVM classifier shows 93.84% accuracy.

3 Dataset used and methodology

The Dataset and methodology used in this study are described in this section.

3.1 Parkinson's speech dataset

The publically available dataset from the UCI repository has been used in this study. The dataset contains speech signals from 8 healthy individuals and 23 PD patients. The dataset was created by Max Little of the University of Oxford. The 195 rows in the dataset correspond to the voice measure of 31 individuals and every column corresponds to a particular voice feature. Out of 195, 147 voice measures are from PD patients and the remaining are from healthy persons. The status column has two values, zero for healthy individuals and one for PD patients. The description of voice features is shown in Table 1 (Little et al., 2008).

3.2 Methodology

The proposed methodology aims to diagnose Parkinson's using speech signals. The proposed hybrid system consists of feature selection, class balancing, and classification stages as shown in Fig. 1. Features from UCI Parkinson's speech data set are selected by three feature selection methods. The feature selection methods used are extra tree, mutual information gain and genetic algorithm. There is a problem of class imbalance in the dataset because out of 195 instances, 147 instances are related to PD patients and the rest related to healthy individuals. Many researchers have ignored this issue but it affects the performance of the classification system. This problem can be resolved by oversampling and undersampling methods. To deal with this issue, synthetic minority oversampling technique (SMOTE) is used in this study. Minority samples in SMOTE are generated synthetically from existing examples in the dataset. The classification was performed naive bayes, k-nearest neighbours and random forest classifiers.

3.3 Feature selection methods

Feature selection is the most popular pre-processing technique widely used to identify important features from a dataset. The advantage of using the optimal feature selection method is twofold; it reduces the dimensionality of data as well as enhancing the performance of the model (Remeseiro & Bolon-Canedo, 2019). Three Feature selection methods are used in this study and 5 and 11 features out of 23 features

Table 1 Features of Parkinson's speech dataset

Feature name	Description
MDVP: Fo(Hz), MDVP: Flo(Hz), MDVP: Fhi(Hz)	Vocal fundamental frequency (Average, Minimum, Maximum)
MDVP: Jitter (%), MDVP: PPQ, MDVP: Jitter (Abs), Jitter: DDP, MDVP: RAP	Fundamental frequency variation measures
MDVP: Shimmer, MDVP: Shimmer(dB), MDVP: APQ, Shimmer: APQ3, Shimmer: APQ5, Shimmer: DDA	Amplitude variation measures
NHR, HNR	Ratio of noise to tonal components measures in the voice
RPDE, D2	Two nonlinear dynamical complexity measures
DFA	Signal fractal scaling exponent
spread1, spread2, PPE	Fundamental frequency variation measures
Status	Health status: (1)—Parkinson's, (0)—Healthy

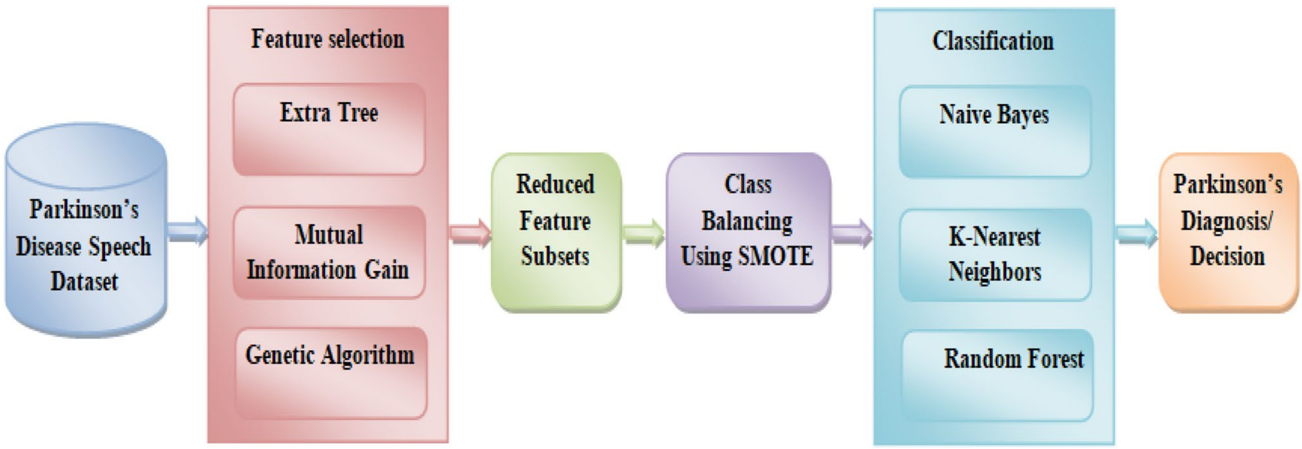

Fig. 1 Methodology of proposed hybrid Parkinson's disease diagnostic system

are selected. The feature selection methods used in the proposed system are described below.

3.3.1 Feature importance using extra tree classifier algorithm

A randomized tree classifier popularly known as the extra tree classifier is an ensemble method in which the feature is selected by using multiple decision trees. The importance of features is calculated using the ensemble of decision trees and features with less importance are removed. In an extra tree classifier, decision trees are constructed from the original training samples. The random samples of k features from the data set are provided to each decision tree at every node and by using a mathematical criterion (Gini Index) best features to split the dataset are selected by every decision tree (Hemphill et al., 2014).

3.3.2 Mutual information gain algorithm

It is a filter method of feature selection. The information gain of each feature is calculated in the context of the dependence of the target feature on the selected feature. The value of information gain ranges from zero to one. Higher information gain features are chosen because they contribute more to the classification process (Remeseiro & Bolon-Canedo, 2019).

3.3.3 Genetic algorithm (GA)

GA is a wrapper method of feature selection that comes in the category of evolutionary algorithms (Murthy & Koolagudi, 2018). To find the best solution, GA works on a population of individuals called chromosomes. From each generation, the best individual is selected by a fitness function and passed to the next generation using a crossover operator. Some random changes are made in the selected individual by a mutation operator for diversity (Chandrashekar & Sahin, 2014).

3.4 Machine learning classifiers

The performance of three classifiers is analyzed on full features and reduced feature subsets. The classifiers used in the proposed methodology is described below.

3.4.1 Naive Bayes (NB)

NB is a probabilistic classification algorithm that works on Bayes theorem to predict the output (Sarker et al., 2019). The researchers have used this classifier in various fields due to its simplicity such as spam filtering, document categorization and disease prediction or diagnosis. This classifier works on the assumption that a particular feature of a given class is not directly related to any other features known as class conditional independence (Murthy & Koolagudi, 2018). In Naive Bayes, the probability of an object having given features belonging to a class called posterior probability P(m|n) is calculated by the equation given below.

$$P(m|n) = P(n|m)P(m) \div P(n)$$

Here P(n|m) denotes the probability of an attribute to be associated with a class. P(m) is Prior class probability. P(n) is the prior probability of attribute.

3.4.2 K nearest neighbor (K-NN)

K-NN is a supervised machine learning classifier that is used for both classification as well as regression. Lazy learning or instance-based learning is used in the k-nn classifier (Sarker et al., 2019). Lazy learning means during classification whole training data is used because there is no particular training phase. The working of k-nn is very simple. Suppose there are two classes in the dataset, a new instance belongs to which class is identified by k-nearest neighbors calculated by Euclidean distance. The class of each neighbor is calculated. The class of new instances is assigned by majority voting between neighbors (Rani et al., 2020; Uddin et al., 2019).

3.4.3 Random forest (RF)

RF is an ensemble supervised machine learning algorithm that can be used for regression as well as classification. The random forest creates many decision trees and randomly chooses the data samples (Rani et al., 2021a). Each decision tree gives the prediction and finally, the best result is chosen by doing voting (Rani et al., 2021b). A new sample is classified by passing down through each decision tree of the forest. Each decision tree is trained by a different part of the dataset so it gives different outcomes. The final classification outcome is done by either most votes (for classification) or the average of all the trees (for regression) in the forest. Figure 2 shows the random forest with three different decision trees. Each tree is trained by random samples of data. Each feature (f1-f9) is represented by circles and the output class (X, Y) is represented by rectangles (Uddin et al., 2019).

3.5 Performance parameters

The performance of classifiers is evaluated in terms of the following parameters.

- Accuracy: Accuracy measures the correct predictions out of total predictions performed.

$$Accuracy = \frac{CorrectPredictions}{TotalPredictions}$$

- Sensitivity: Sensitivity also known as True Positive Rate (TPR) or Recall measures the ability of a system to make correct positive predictions.

$$TPR = \frac{TruePostives}{TruePositives + FalseNegatives}$$

- Specificity: Specificity also known as True Negative Rate (TNR) measures the ability of a system to make correct negative predictions.

$$TNR = \frac{TrueNagatives}{TrueNegatives + FalsePositives}$$

- Precision: Precision also known as Positive Prediction Value (PPV) measures the capability of a system to produce only relevant results.

$$Precision = \frac{TruePositives}{TruePositives + FalsePositives}$$

- F-Measure: F-Measure calculates the harmonic mean of precision and recall.

$$F - Measure = 2 * \frac{Recall * Precision}{Recall + Precision}$$

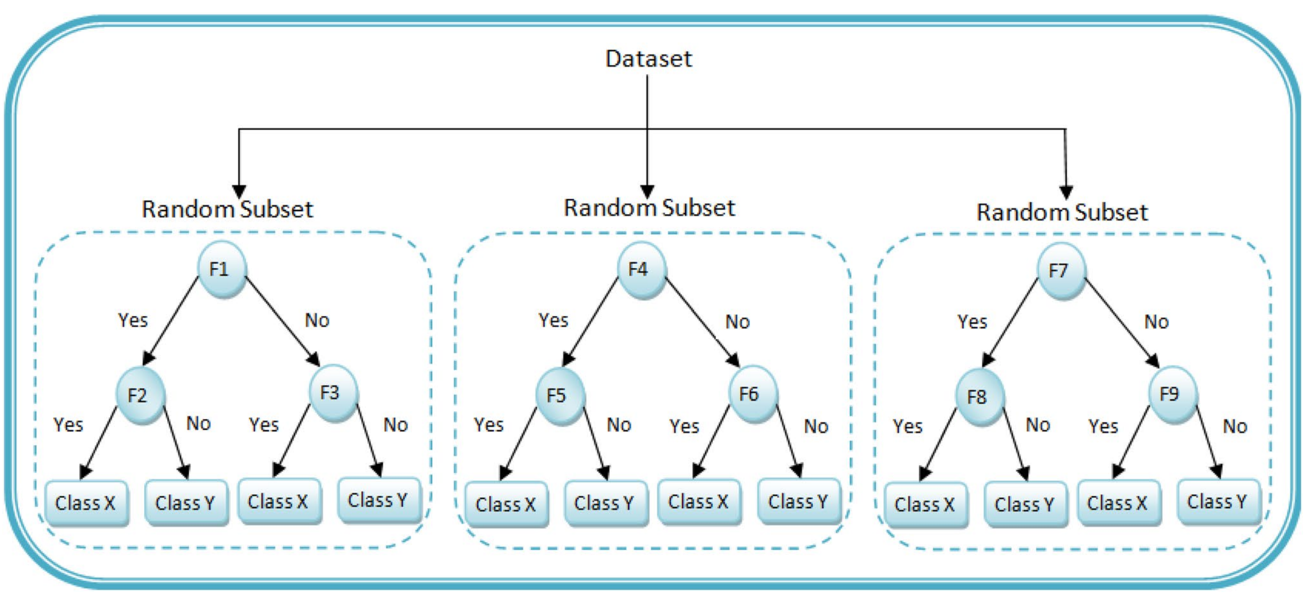

Fig. 2 Random Forest with three different decision trees

4 Experimental results and discussions

This section gives a comparative analysis of feature selection methods with different classifiers. Three feature selection methods extra tree, mutual information gain, and genetic algorithm, and three classifiers naïve bayes, k- nearest neighbors, and random forest are used in this study. k-fold cross-validation method with k = 10 is used for validating the results. Accuracy, specificity, f-measure, sensitivity, precision, and AUC are used as performance measures. The features having high importance score perform well during the classification process. The feature importance of each feature is calculated by the extra tree method, and then the top 11 features are selected. In mutual information gain method, information gain of each feature is calculated. The features having high information gain value are used for the classification process. 11 features are selected by this method. The genetic algorithm performs well among three feature selection methods and selects only 5 features. Table 2 shows the features selected by feature selection methods.

After selecting the features by different feature selection methods, the performance of classifiers were evaluated on different feature subsets. The performance of the naïve bayes classifier on all features and reduced feature subsets selected by feature selection methods is shown in Table 3. The genetic algorithm outperforms among three feature selection methods with 84.67% accuracy.

The performance of the k-nearest neighbor classifier on all features and reduced feature subsets selected by feature selection methods is shown in Table 4. The genetic algorithm outperforms among three feature selection methods with 91.45% accuracy.

The performance of the random forest classifier on all features and reduced feature subsets selected by feature selection methods is shown in Table 5. The genetic algorithm outperforms among three feature selection methods with 95.58% accuracy.

A comparative analysis of improvement in accuracy of classifiers by feature selection method is shown in Fig. 3. The accuracy of all three classifiers is improved by using reduced feature subsets selected by feature selection methods. Genetic algorithms have contributed more to improve the performance of all classifiers than other feature selection methods. The performance of the naïve bayes classifier is improved by 21.70%, the performance of k-nearest neighbor is improved by 20.83% and the performance of the random forest algorithm is improved by 12.52%. The

Table 2 Features selected by feature selection methods

Feature selection method	No of feature selected	Features
Extra tree	11	MDVP:Fo(Hz), MDVP:Flo(Hz), MDVP:Fhi(Hz), MDVP:Jitter(Abs), Shimmer:APQ5, DFA, NHR, PPE, RPDE, spread1, spread2
Mutual information gain algorithm	11	MDVP:Fo(Hz), MDVP:Flo(Hz), MDVP:Fhi(Hz), MDVP:Jitter(Abs), Shimmer:APQ5, MDVP:Shimmer(dB), MDVP:APQ, HNR, PPE, spread1, spread2
Genetic algorithm	05	MDVP:Fo(Hz), MDVP:Shimmer(dB), MDVP:PPQ, spread1, spread2

Table 3 Performance of naïve bayes classifier with feature selection methods

Feature selection algorithm	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	F-measure (%)	AUC
All features	69.57	63.94	87.50	94.00	76.11	0.820
Extra tree	80.63	76.19	85.03	83.58	79.71	0.898
Mutual information gain	84.00	73.46	94.55	93.10	82.12	0.903
Genetic algorithm	84.67	73.46	95.91	94.73	82.75	0.912

Table 4 Performance of k-nearest neighbor with feature selection methods

Feature selection algorithm	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	F-measure (%)	AUC
All features	75.68	82.99	54.16	84.72	83.48	0.686
Extra tree	86.35	78.91	93.87	92.80	85.29	0.864
Mutual information gain	86.37	77.55	95.23	94.21	85.07	0.864
Genetic algorithm	91.45	89.11	93.87	93.57	91.28	0.915

combination of the genetic algorithm and random forest classifier gives the best accuracy of 95.58%.

Receiver operating characteristic curve (ROC) has been also used as a performance parameter. ROC is a graphical plot by which the performance of the classifier is presented. ROC is plotted between the true-positive rate and false-positive rate. The value of AUC is ranging between zero and one. The higher the values of AUC, the system perform better in differentiating between patients with healthy individuals. The performance of classifiers on features selected by the genetic algorithm is shown in Fig. 4. All the classifiers have performed well in terms of AUC

Table 5 Performance of random forest classifier with feature selection methods

Feature selection algorithm	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	F-Measure (%)	AUC
All features	84.94	94.55	56.25	86.87	90.55	0.896
Extra tree	94.58	91.83	97.27	97.12	94.40	0.987
Mutual information gain	94.57	93.19	95.91	95.80	94.48	0.987
Genetic algorithm	95.58	93.19	97.95	97.85	95.47	0.989

Fig. 3 Comparative analysis of feature selection methods with classifiers

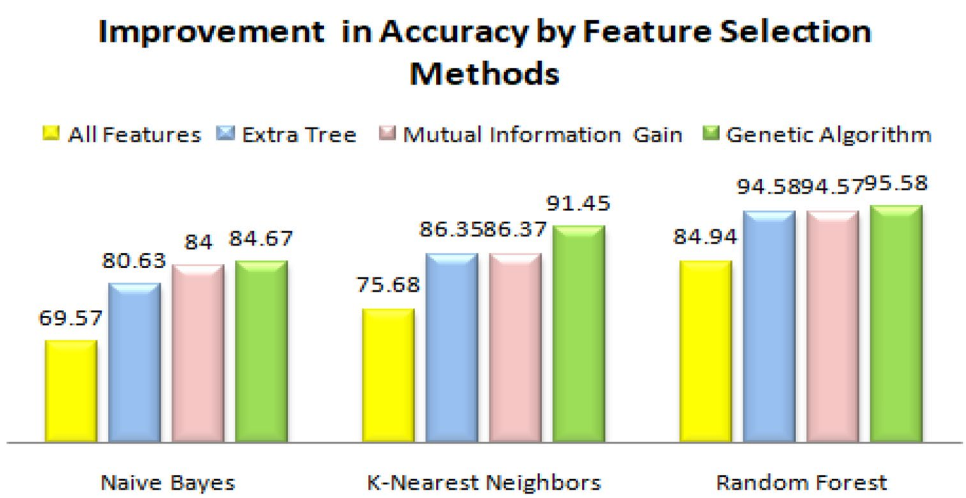

Fig. 4 ROC Plot

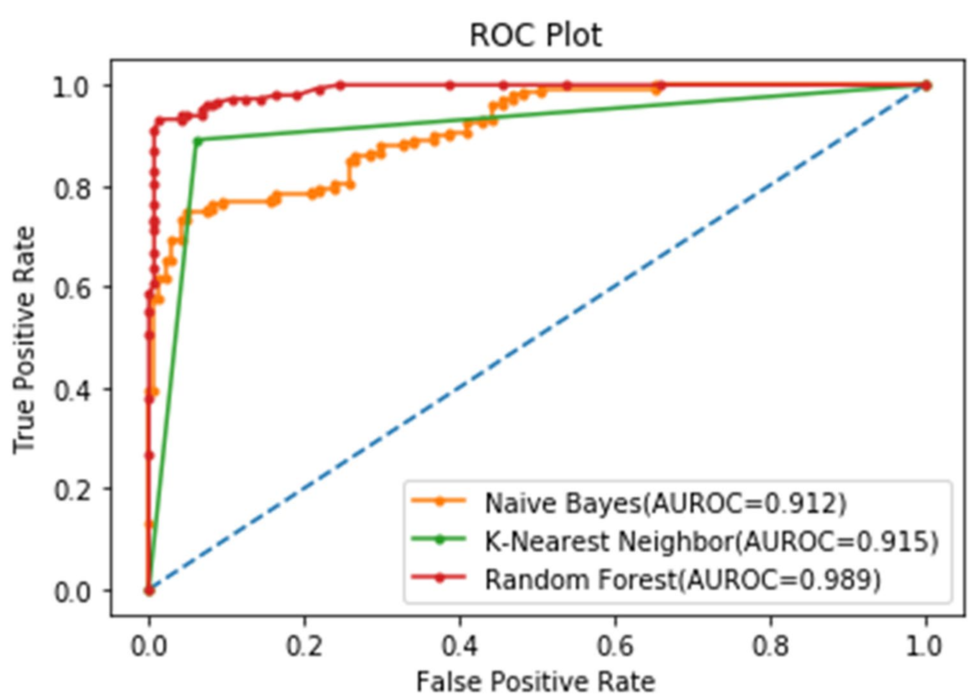

Table 6 Comparison with the existing system

Study	Dataset used	Feature selection methods	Classifies	Accuracy (%)
Gupta et al. (2018)	UCI Parkinson's Speech Dataset	Optimized Cuttlefish Algorithm	K-Nearest Neighbor, Decision Tree	92.19
Sharma, Sundaram, et al. (2019)	UCI Parkinson's Speech Dataset	Modified Grey Wolf Optimization	K-Nearest Neighbor, Random Forest, Decision Tree	93.87
Senturk (2020)	UCI Parkinson's Speech Dataset	Recursive Feature Elimination, Feature Importance	Support Vector Machines, Artificial Neural Networks and Classification and Regression Trees	93.84
Proposed hybrid system	UCI Parkinson's Speech Dataset	Extra Tree, Mutual Information Gain, Genetic Algorithm	Naïve Bayes, K-Nearest Neighbor, Random Forest	95.58

and the random forest classifier achieved the highest AUC 0.989.

Proposed hybrid Parkinson's disease diagnosis system performed better than the existing systems. The comparative analysis is shown in Table 6. The accuracy enhancement of the proposed hybrid method compared to existing systems is shown in Fig. 5.

5 Conclusion and future scope

Parkinson's is a chronic disease hence, the only way to improve a patient's life is early detection. Proper medication, a balanced diet and exercise can lower the symptoms of Parkinson's. This study is focused on the early diagnosis of Parkinson's disease by speech signals. There is no universal feature selection method and classifier for the medical dataset. For finding the best results one has to try different methods. Authors have tried different methods and found the best combination. The results suggest that the use of the feature selection method is advantageous because it reduces complexity and increases accuracy. The proposed hybrid system outperforms some recent existing works with 95.58% accuracy. The proposed system is not a replacement for health care professionals but can be used as a second opinion for the Parkinson diagnosis. In the future, authors have plans to test the proposed method on significant speech and voice datasets. Authors also plan to make Parkinson's disease diagnosis system by handwritten drawings as slowness and tremors is also an early symptom of Parkinson's which adversely affects the patient's handwriting.

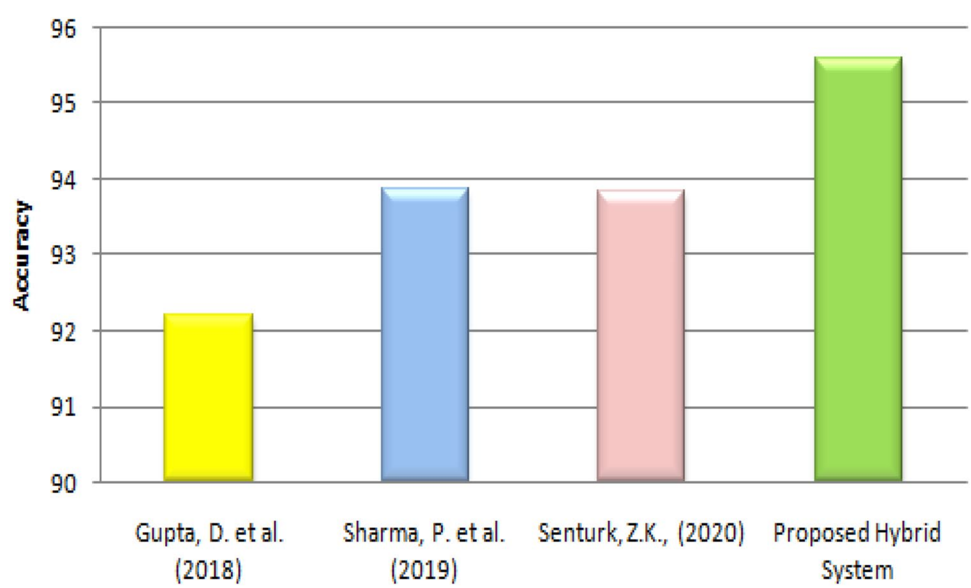

Fig. 5 Comparison with existing systems

Declarations

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

References

Almeida, J. S., Rebouças Filho, P. P., Carneiro, T., Wei, W., Damaševičius, R., Maskeliūnas, R., & de Albuquerque, V. H. C. (2019). Detecting Parkinson's disease with sustained phonation and speech signals using machine learning techniques. *Pattern Recognition Letters, 125*, 55–62. https://doi.org/10.1016/j.patrec.2019.04.005.

Chandrashekar, G., & Sahin, F. (2014). A survey on feature selection methods. *Computers & Electrical Engineering, 40*(1), 16–28. https://doi.org/10.1016/j.compeleceng.2013.11.024.

Emamzadeh, F. N., & Surguchov, A. (2018). Parkinson's disease: Biomarkers, treatment, and risk factors. *Frontiers in Neuroscience, 12*, 612. https://doi.org/10.3389/fnins.2018.00612.

Grover, S., Bhartia, S., Yadav, A., & Seeja, K. R. (2018). Predicting severity of Parkinson's disease using deep learning. *Procedia Computer Science, 132*, 1788–1794. https://doi.org/10.1016/j.procs.2018.05.154.

Gunduz, H. (2019). Deep learning-based Parkinson's disease classification using vocal feature sets. *IEEE Access, 7*, 115540–115551. https://doi.org/10.1109/ACCESS.2019.2936564.

Gupta, D., Julka, A., Jain, S., Aggarwal, T., Khanna, A., Arunkumar, N., & de Albuquerque, V. H. C. (2018). Optimized cuttlefish algorithm for diagnosis of Parkinson's disease. *Cognitive Systems Research, 52*, 36–48. https://doi.org/10.1016/j.cogsys.2018.06.006.

Hemphill, E., Lindsay, J., Lee, C., Măndoiu, I. I., & Nelson, C. E. (2014). Feature selection and classifier performance on diverse bio-logical datasets. *BMC bioinformatics, BioMed Central, 15*(13), S4. https://doi.org/10.1186/1471-2105-15-S13-S4.

Jain, A., Tiwari, S., & Sapra, V. (2019). Two-phase heart disease diagnosis system using deep learning. *International Journal of Control and Automation*, *12*(5), 558–573.

Lahmiri, S., & Shmuel, A. (2019). Detection of Parkinson's disease based on voice patterns ranking and optimized support vector machine. *Biomedical Signal Processing and Control, 49*, 427–433. https://doi.org/10.1016/j.bspc.2018.08.029.

Lamba, R., Gulati, T., Al-Dhlan, K. A., & Jain, A. (2021). A systematic approach to diagnose Parkinson's disease through kinematic features extracted from handwritten drawings. *Journal of Reliable Intelligent Environments*, 1–10. https://doi.org/10.1007/s40860-021-00130-9.

Lamba, R., Gulati, T., & Jain, A. (2020). Comparative analysis of Parkinson's disease diagnosis system. *Advances in Mathematics Scientific Journal, 9*(6), 3399–3406. https://doi.org/10.37418/amsj.9.6.20.

Little, M., McSharry, P., Hunter, E., Spielman, J., & Ramig, L. (2008). Suitability of dysphonia measurements for telemonitoring of Parkinson's disease. *Nature Precedings*. https://doi.org/10.1038/npre.2008.2298.1.

Mostafa, S. A., Mustapha, A., Mohammed, M. A., Hamed, R. I., Arunkumar, N., Ghani, M. K. A., Jaber, M. M., & Khaleefah, S. H. (2019). Examining multiple feature evaluation and classification methods for improving the diagnosis of Parkinson's disease. *Cognitive Systems Research, 54*, 90–99. https://doi.org/10.1016/j.cogsys.2018.12.004.

Murthy, Y. S., & Koolagudi, S. G. (2018). Classification of vocal and non-vocal segments in audio clips using genetic algorithm based feature selection (GAFS). *Expert Systems with Applications, 106*, 77–91. https://doi.org/10.1016/j.eswa.2018.04.005.

Naranjo, L., Pérez, C. J., Campos-Roca, Y., & Martín, J. (2016). Addressing voice recording replications for Parkinson's disease detection. *Expert Systems with Applications, 46*, 286–292. https://doi.org/10.1016/j.eswa.2015.10.034.

Nissar, I., Rizvi, D., Masood, S. and Mir, A. (2019). Voice-Based Detection of Parkinson's Disease through Ensemble Machine Learning Approach: A Performance Study. *EAI Endorsed Transactions on Pervasive Health and Technology*, 5(19). doi: https://doi.org/10.4108/eai.13-7-2018.162806

Olivares, R., Munoz, R., Soto, R., Crawford, B., Cárdenas, D., Ponce, A., & Taramasco, C. (2020). An optimized brain-based algorithm for classifying Parkinson's disease. *Applied Sciences, 10*(5), 1827. https://doi.org/10.3390/app10051827.

Polat, K. (2019). A hybrid approach to Parkinson disease classification using speech signal: The combination of SMOTE and random forests. *In 2019 scientific meeting on electrical-electronics & biomedical engineering and computer science (EBBT) IEEE,* 1–3. https://doi.org/10.1109/EBBT.2019.8741725

Rani, P., Kumar, R., Ahmed, N. M. O. S., & Jain, A. (2021a). A decision support system for heart disease prediction based upon machine learning. *Journal of Reliable Intelligent Environments*. https://doi.org/10.1007/s40860-021-00133-6.

Rani, P., Kumar, R., & Jain, A. (2021b). Multistage Model for Accurate Prediction of Missing Values Using Imputation Methods in Heart Disease Dataset. In J. S. Raj, A. M. Iliyasu, R. Bestak, & Z. A. Baig (Eds.), *Innovative Data Communication Technologies and Application, Lecture Notes on Data Engineering and Communications Technologies* (pp. 637–653). Singapore: Springer. https://doi.org/10.1007/978-981-15-9651-3_53.

Rani, P., Kumar, R., Jain, A., & Lamba, R. (2020). Taxonomy of Machine Learning Algorithms and its Applications. *Journal of Computational and Theroretical Nanoscience, 17*(6), 2509–2514. https://doi.org/10.1166/jctn.2020.8922.

Reich, S. G., & Savitt, J. M. (2018). Parkinson Disease. *Medical Clinics of North America*. https://doi.org/10.1016/j.mcna.2018.10.014.

Remeseiro, B., & Bolon-Canedo, V. (2019). A review of feature selection methods in medical applications. *Computers in Biology and Medicine*. https://doi.org/10.1016/j.compbiomed.2019.103375.

Sakar, C. O., Serbes, G., Gunduz, A., Tunc, H. C., Nizam, H., Sakar, B. E., Tutuncu, M., Aydin, T., Isenkul, M. E., & Apaydin, H. (2019). A comparative analysis of speech signal processing algorithms for Parkinson's disease classification and the use of the tunable Q-factor wavelet transform. *Applied Soft Computing, 74*, 255–263. https://doi.org/10.1016/j.asoc.2018.10.022.

Sarker, I. H., Kayes, A. S. M., & Watters, P. (2019). Effectiveness analysis of machine learning classification models for predicting personalized context-aware smartphone usage. *Journal of Big Data, 6*(1), 57. https://doi.org/10.1186/s40537-019-0219-y.

Senturk, Z. K. (2020). Early diagnosis of Parkinson's disease using machine learning algorithms. *Medical Hypotheses, 138*, 109603. https://doi.org/10.1016/j.mehy.2020.109603.

Sharma, P., Jain, R., Sharma, M., & Gupta, D. (2019). Parkinson's diagnosis using ant-lion optimisation algorithm. *International Journal of Innovative Computing and Applications, 10*(3–4), 138–146. https://doi.org/10.1504/IJICA.2019.103370.

Sharma, P., Sundaram, S., Sharma, M., Sharma, A., & Gupta, D. (2019). Diagnosis of Parkinson's disease using modified grey wolf optimization. *Cognitive Systems Research, 54*, 100–115. https://doi.org/10.1016/j.cogsys.2018.12.002.

Solana-Lavalle, G., Galán-Hernández, J. C., & Rosas-Romero, R. (2020). Automatic Parkinson disease detection at early stages as a pre-diagnosis tool by using classifiers and a small set of vocal features. *Biocybernetics and Biomedical Engineering, 40*(1), 505–516. https://doi.org/10.1016/j.bbe.2020.01.003.

Tuncer, T., & Dogan, S. (2019). A novel octopus based Parkinson's disease and gender recognition method using vowels. *Applied Acoustics, 155*, 75–83. https://doi.org/10.1016/j.apacoust.2019.05.019.

Tuncer, T., Dogan, S., & Acharya, U. R. (2020). Automated detection of Parkinson's disease using minimum average maximum tree and singular value decomposition method with vowels. *Biocybernetics and Biomedical Engineering, 40*(1), 211–220. https://doi.org/10.1016/j.bbe.2019.05.006.

Uddin, S., Khan, A., Hossain, M. E., & Moni, M. A. (2019). Comparing different supervised machine learning algorithms for disease prediction. *BMC Medical Informatics and Decision Making, 19*(1), 1–16. https://doi.org/10.1186/s12911-019-1004-8.

Xiong, Y., & Lu, Y. (2020). Deep Feature Extraction From the Vocal Vectors Using Sparse Autoencoders for Parkinson's Classification. *IEEE Access, 8*, 27821–27830. https://doi.org/10.1109/ACCESS.2020.2968177.

Yaman, O., Ertam, F., & Tuncer, T. (2020). Automated Parkinson's disease recognition based on statistical pooling method using acoustic features. *Medical Hypotheses, 135*, 109483. https://doi.org/10.1016/j.mehy.2019.109483.

Zahid, L., Maqsood, M., Durrani, M. Y., Bakhtyar, M., Baber, J., Jamal, H., Mehmood, I., & Song, O. Y. (2020). A Spectrogram-Based Deep Feature Assisted Computer-Aided Diagnostic System for Parkinson's Disease. *IEEE Access, 8*, 35482–35495. https://doi.org/10.1109/ACCESS.2020.2974008.

Zesiewicz, T. A., Bezchlibnyk, Y., Dohse, N., & Ghanekar, S. D. (2019). Management of Early Parkinson Disease. *Clinics in geriatric medicine*. https://doi.org/10.1016/j.cger.2019.09.001.