Available online at www.medicinescience.org

ORIGINAL ARTICLE

Medicine Science International Medical Journal

Medicine Science 2025;14(2):355-61

Analysis of voice data based on associative classification rules for early diagnosis of Parkinson's disease

Ipek Balikci Cicek, Zeynep Kucukakcali

İnönü University, Faculty of Medicine, Department of Biostatistics and Medical Informatics, Malatya, Türkiye

Received 24 October 2024; Accepted 24 January 2025
Available online 30 May 2025 with doi: 10.5455/medscience.2024.10.130

Abstract

Parkinson's disease (PD) is a neurodegenerative disorder characterized by the gradual degradation of dopaminergic neurons in the substantia nigra pars compacta. The condition presents motor symptoms (such as tremors, rigidity, and bradykinesia) alongside non-motor symptoms (including cognitive impairments and sleep disturbances). Early diagnosis of PD is challenging since symptoms typically appear at advanced stages. In this context, novel methods such as voice analysis have gained importance in detecting the disease in its early stages. Alterations in vocal attributes in persons with PD manifest early in the condition and employing machine learning algorithms to assess these vocal qualities can serve as an effective instrument for early detection. This study aims to examine the utilization of speech analysis for the early detection of PD and the feasibility of machine learning algorithms in these assessments. The study utilizes a PD speech dataset sourced from the UCI Machine Learning repository, comprising 195 voice recordings from 31 individuals, 23 of whom are diagnosed with PD. The study's principal objective is to distinguish between healthy persons and those with PD based on vocal attributes. The associative classification approach was utilized to derive associations between voice characteristics as "If-X then Y" rules. The Ameva algorithm was used to generate the rules evaluated based on support and confidence metrics, selecting the most significant ones for classification. The classification algorithm achieved an accuracy of 92.8%, a sensitivity of 90.5%, and a specificity of 100%. These results indicate that certain voice features, such as fundamental frequency, shimmer, spread, and PPE, are strong indicators for diagnosing PD. The association rules, which showed a perfect confidence level, highlight the importance of these voice parameters in accurately identifying individuals with PD. This study demonstrates the potential of voice analysis, using associative classification, for the early diagnosis of PD. The model showed high accuracy and sensitivity, making voice features like fundamental frequency and shimmer effective indicators of PD. The use of "if-then" rules enhances the interpretability and practical application of the method. Overall, this approach could significantly aid in the early detection of Parkinson's, improving patient outcomes through timely interventions.

Keywords: Parkinson's disease, associative classification, machine learning, voice analysis, early diagnosis

Introduction

Parkinson's disease (PD) is a neurological condition that results in the gradual loss of dopaminergic neurons in the substantia nigra pars compacta (SNc) [1]. The illness usually affects individuals between the ages of 70 and 79, with the highest incidence found in those aged 85 to 89 [2]. Symptoms of the illness include tremors, stiffness, and poor balance and coordination [3]. The incidence of PD is around 1% in those over 60, affecting about 1 to 2 individuals per 1,000 [4]. It is considered one of the most common neurodegenerative illnesses, with a worldwide prevalence of 100-200 cases per 100,000 individuals [5]. Current demographic trends indicate that the incidence of PD is expected to quadruple by 2030, with a lifetime risk approaching 4% [6]. PD is a neurological disorder characterized by both motor and nonmotor symptoms. Tremors, bradykinesia, rigidity, and postural instability exemplify motor symptoms. Non-motor symptoms including cognitive impairments, constipation, sleep disturbances, genitourinary dysfunction, and mood problems [7,8].

PD is one of the most formidable diagnoses in neurology, as it is typically identified at an advanced stage [9]. Braak et al. [10] delineate six neuropathological phases that manifest as the disease advances. Olfactory and vocal aberrations manifest throughout the initial two phases. Motor symptoms become increasingly apparent

CITATION
Balikci Cicek I, Kucukakcali Z. Analysis of voice data based on associative classification rules for early diagnosis of Parkinson's disease. Med Science. 2025;14(2):355-61.

Corresponding Author: Ipek Balikci Cicek, İnönü University, Faculty of Medicine, Department of Biostatistics and Medical Informatics, Malatya, Türkiye
Email: ipek.balikci@inonu.edu.tr

DOI: 10.5455/medscience.2024.10.130

in the third and fourth phases, facilitating a physician's diagnosis of PD. Extensive regions of the brain undergo significant alterations throughout the fifth and sixth phases. Dysphonia, hypophonia, monotone pitch range, and dysarthria are vocal abnormalities observed in persons with PD [11].

Although PD is incurable, early diagnosis and treatment may enhance patients' quality of life. Vocal irregularities are prodromal symptoms that arise in 90 percent of Parkinson's patients in their early stages [12]. Thus, research into voice features has been used to detect PD in its early stages. Furthermore, the voice analysis technique is inexpensive and easy to use at first. Using pattern recognition on speech recordings as an aid in PD diagnosis has various advantages [13]. To begin, voice abnormalities can be identified early in the disease's progression. Second, voice recordings are easy to get. Third, there is a diverse set of signal processing algorithms accessible for extracting speech information. Fourth, applying machine learning techniques to classify voice characteristics can help with illness detection early on. Finally, early detection of PD and effective treatment might potentially improve patients' quality of life. Using machine learning algorithms to categorize vocal characteristics aids in the early detection of illness.

Machine learning (ML) is increasingly being used to diagnose medical conditions due to its ease of use and high accuracy [14]. In fact, the literature notes its application in treating PD as well [15]. Associative classification (AC) is a unique type of machine learning technique that employs association rule mining for classification. Several studies have shown that AC yields more accurate results compared to other traditional classification techniques in data mining [16]. As a subfield of machine learning, AC categorizes unseen instances by leveraging association rules. It has been extensively studied over the past two decades, particularly in the fields of text classification and medical diagnostics. The AC method offers several advantages. First, it achieves higher classification accuracy compared to alternative methods. Additionally, the rules generated by AC classifiers are simple and presented in easy-to-understand "if-then" statements, enabling users to modify and manage them as needed. The main goal of an AC strategy is to construct a classifier (model) from a large dataset (training data) to predict or identify the category of new instances (test data) [17].

The aim of this study is to apply the associative classification method, a machine learning model, to enhance the early and accurate diagnosis of PD. By identifying the disease at earlier stages, this approach not only facilitates timely intervention but also improves patient outcomes, ultimately aiding physicians in delivering more effective treatments and enhancing the overall quality of life for individuals with PD.

Material and Methods

Dataset

The PD speech dataset was obtained from the UCI Machine Learning repository located at the Centre for Machine Learning and Intelligent Systems [18]. This dataset comprises several biological speech metrics obtained from 31 individuals, 23 of whom have been diagnosed with PD. Each column in the table denotes a particular vocal measurement, while each row corresponds to one of the 195 voice recordings from the individuals, as specified in the "name" column. The main aim of the data is to distinguish persons in good health from those with PD, as shown by the "status" column, where a value of 0 signifies good health and a value of 1 denotes PD.

The dataset contains a list of variables, including their names, kinds, and descriptions, which are presented in Table 1.

Table 1. Descriptions of variables in the Parkinson's disease dataset

Variable name	Description
Name	ASCII subject name and recording number
MDVP: Fhi (Hz)	Maximum vocal fundamental frequency
MDVP: Fo (Hz)	Average vocal fundamental frequency
MDVP: Jitter (%), MDVP: Jitter (Abs), MDVP: RAP, MDVP: PPQ, Jitter: DDP	Several measures of variation in fundamental frequency
MDVP:Flo(Hz)	Minimum vocal fundamental frequency
MDVP: Shimmer, MDVP: Shimmer (dB), Shimmer: APQ3, Shimmer: APQ5, MDVP: APQ, Shimmer: DDA	Several measures of variation in amplitude
RPDE, D2	Two nonlinear dynamical complexity measures
NHR, HNR	Two measures of the ratio of noise to tonal components in the voice
Spread1, Spread2, PPE	Three nonlinear measures of fundamental frequency variation
DFA	Signal fractal scaling exponent
Status: The health status of the subject	(one): Parkinson's, (zero): healthy

DOI: 10.5455/medscience.2024.10.130

Associative classification Method

Associative classification is a machine learning method that merges association rule mining and classification techniques to build models that uncover strong relationships between data attributes. AC is particularly suitable for applications requiring transparency and strong association discovery, such as market basket analysis, medical diagnosis, fraud detection, customer segmentation, and bioinformatics. The method's flexibility allows it to handle various types of data, making it applicable across different domains and data types.

This method begins with data preparation, where the dataset is structured with categorical attributes, and continuous attributes are discretized if necessary. Association rule mining algorithms such as Apriori or FP-Growth are then used to generate rules in the form "If X then Y," where X represents a set of attributes and Y represents the class label. These rules are evaluated based on support and confidence metrics, which indicate the frequency and likelihood of the rules being correct, respectively [19,20]. This method offers several advantages, including high interpretability and the ability to uncover strong feature associations. However, it also faces challenges related to computational complexity, redundancy, and overfitting. To address these challenges, enhanced pruning techniques, hybrid approaches, and optimization algorithms are implemented [21,22]. The Ameva method enhances this associative classification approach by providing an optimal discretization technique specifically designed for improving classification accuracy. Within the associative classification framework, Ameva automatically generates membership functions that define precise boundaries between categories. This integration is particularly valuable for medical datasets where accurate class distinctions are crucial. The method's ability to generate optimal cut points for continuous variables strengthens the quality of the association rules generated, leading to more reliable classification outcomes. Furthermore, Ameva's discretization process aligns with associative classification's need for categorical attributes, making it an ideal preprocessing step in the overall classification pipeline [23].

Model Evaluation

The model was evaluated using various performance metrics, including accuracy, sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and F1-score. In this study, all analyses were performed using the R programming language. The arules package was used to create and evaluate association rules in data analysis, the caret package was used for modeling and calculating performance metrics, the AmevaR package was used to apply the Ameva algorithm for discretization of continuous variables, and the ggplot2 package was used to visualize graphical results.

Results

Table 2 provides the values for the model's classification performance metrics, including accuracy, balanced accuracy, sensitivity, specificity, positive and negative predictive values, and the F1-score.

Table 2. Values of the model's classification performance metrics

Metric	Value
Accuracy	0.928
Balanced accuracy	0.952
Sensitivity	0.905
Specificity	1
Positive predictive value	1
Negative predictive value	0.774
F1-score	0.95

These metrics represent the performance of a classification model. Here's the interpretation of each metric:

1. **Accuracy – 0.928:** The overall accuracy of the model is 92.8%. This means the model correctly classified 92.8% of the instances across all data.
2. **Balanced Accuracy – 0.952:** Balanced accuracy takes into account any imbalance between the classes and is 95.2%. This indicates the model performed well on both classes.
3. **Sensitivity – 0.905:** The sensitivity, or recall, is 90.5%, which means the model correctly identified 90.5% of the actual positive cases. Although the model is good at detecting positives, it may still miss some true positives.
4. **Specificity – 1:** The model's specificity is 100%, meaning it perfectly identified all the negative cases. There are no false positives.
5. **Positive Predictive Value – 1:** The model has a perfect positive predictive value of 100%. This indicates that all instances predicted as positive are indeed positive, with no false positives.
6. **Negative Predictive Value – 0.774:** The negative predictive value is 77.4%, meaning 77.4% of the cases predicted as negative are truly negative. This suggests the model is less reliable in its negative predictions, with a higher proportion of false negatives.
7. **F1-score – 0.95:** The F1-score, which combines precision and recall, is 95%. This high score shows that the model performs well in identifying positive cases while maintaining precision.

Overall, the model shows strong performance in detecting positive cases (high F1-score), with perfect precision (no false positives). However, it has a somewhat lower reliability in negative predictions (lower NPV), indicating some room for improvement in reducing false negatives.

The classification algorithm's association rules are presented in Table 3.

DOI: 10.5455/medscience.2024.10.130

Table 3. Association rules employed by the classification algorithm

	Left-hand side rules	Right-hand side rules	Support	Confidence	Frequency
Rule 1	{mdvp.fo.hz.=[88.3,195), mdvp.shimmer.db.=[0.198,1.3), spread1=[-6.65,-2.43), ppe=[0.134,0.527)}	{status=1}	0.467	1	91
Rule 2	{mdvp.fhi.hz.=[102,234), mdvp.shimmer.db.=[0.198,1.3), spread1=[-6.65,-2.43), ppe=[0.134,0.527)}	{status=1}	0.456	1	89
Rule 3	{mdvp.fo.hz.=[88.3,195), shimmer.apq5=[0.0132,0.0794), ppe=[0.134,0.527)}	{status=1}	0.451	1	88
Rule 4	{mdvp.fo.hz.=[88.3,195), mdvp.rap=[0.00177,0.0214), dfa=[0.684,0.825)}	{status=1}	0.451	1	88
Rule 5	{mdvp.shimmer.db.=[0.198,1.3), spread1=[-6.65,-2.43),spread2=[0.179,0.45)}	{status=1}	0.426	1	83
Rule 6	{mdvp.rap=[0.00177,0.0214), dfa=[0.684,0.825), spread2=[0.179,0.45)}	{status=1}	0.405	1	79
Rule 7	{mdvp.fo.hz.=[88.3,195),hnr=[8.44,24), dfa=[0.684,0.825),ppe=[0.134,0.527)}	{status=1}	0.4	1	78
Rule 8	{mdvp.fo.hz.=[88.3,195),d2=[2.33,3.67),ppe=[0.134,0.527)}	{status=1}	0.374	1	73
Rule 9	{mdvp.fhi.hz.=[102,234), spread1=[-6.65,-2.43),spread2=[0.179,0.45),d2=[2.33,3.67)}	{status=1}	0.374	1	73
Rule 10	{rpde=[0.47,0.685),dfa=[0.684,0.825), ppe=[0.134,0.527)}	{status=1}	0.354	1	69
Rule 11	{}	{status=0}	0.246	0.246	195

Rule 1 Comment: This rule suggests that individuals with a fundamental frequency between 88.3 and 195 Hz, shimmer (dB) between 0.198 and 1.3, spread1 between -6.65 and -2.43, and PPE between 0.134 and 0.527 are highly associated with PD, as the confidence level is 1. This indicates a perfect correlation between these vocal parameters and the presence of Parkinson's, making them key features for identifying the disease.

Rule 2 Comment: This rule indicates that individuals with a maximum fundamental frequency between 102 and 234 Hz, shimmer (dB) between 0.198 and 1.3, spread1 between -6.65 and -2.43, and PPE between 0.134 and 0.527 are perfectly associated with PD. The high confidence and support suggest that these vocal characteristics are strong indicators of the disease.

Rule 3 Comment: Individuals with a fundamental frequency between 88.3 and 195 Hz, shimmer (APQ5) between 0.0132 and 0.0794, and PPE between 0.134 and 0.527 are perfectly associated with PD, reinforcing the importance of these vocal features in the classification.

Rule 4 Comment: This rule indicates that individuals with a fundamental frequency between 88.3 and 195 Hz, RAP values between 0.00177 and 0.0214, and DFA between 0.684 and 0.825 are strongly correlated with PD, with a perfect confidence level, showing the high reliability of these features in detecting the condition.

Rule 5 Comment: This rule highlights that individuals with shimmer (dB) between 0.198 and 1.3, spread1 between -6.65 and -2.43, and spread2 between 0.179 and 0.45 are perfectly associated with PD, indicating the significance of amplitude and frequency variation in the classification.

DOI: 10.5455/medscience.2024.10.130

Rule 6 Comment: This rule suggests that subjects with RAP values between 0.00177 and 0.0214, DFA between 0.684 and 0.825, and spread2 between 0.179 and 0.45 have a perfect association with PD, indicating that these dynamical complexity and frequency measures are strong predictors of the condition.

Rule 7 Comment: This rule shows that individuals with a fundamental frequency between 88.3 and 195 Hz, HNR between 8.44 and 24, DFA between 0.684 and 0.825, and PPE between 0.134 and 0.527 are perfectly associated with PD, highlighting the combined importance of these vocal and dynamical features.

Rule 8 Comment: This rule suggests that a fundamental frequency between 88.3 and 195 Hz, D2 values between 2.33 and 3.67, and PPE between 0.134 and 0.527 are perfectly associated with PD, showing that these nonlinear dynamical features are crucial in distinguishing patients with the disease.

Rule 9 Comment: This rule indicates that subjects with a maximum fundamental frequency between 102 and 234 Hz, spread1 between -6.65 and -2.43, spread2 between 0.179 and 0.45, and D2 between 2.33 and 3.67 are perfectly associated with PD, showing that these vocal and nonlinear features are significant in the classification process.

Rule 10 Comment: This rule suggests that RPDE values between 0.47 and 0.685, DFA between 0.684 and 0.825, and PPE between 0.134 and 0.527 are perfectly associated with PD, demonstrating the relevance of these nonlinear and fractal scaling features in identifying the disease.

Figure 1 presents the network graph for association rules.

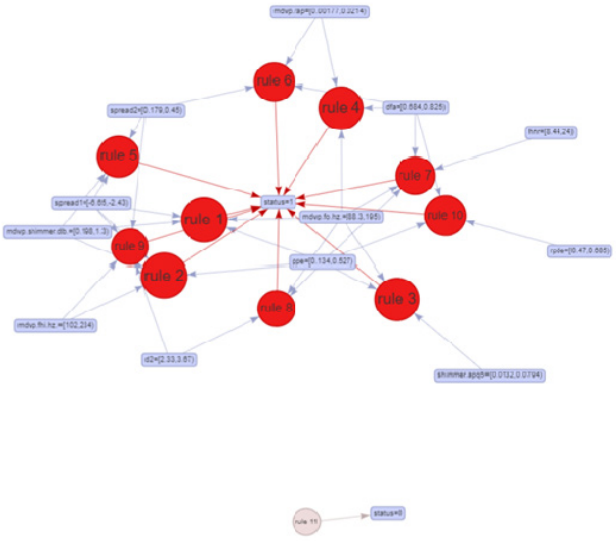

Figure 1. Network graph for association rules

Comment for Figure 1: Network Graph for Association Rules

The network graph in Figure 1 visually represents the association rules applied in the classification algorithm, with a focus on distinguishing between subjects with PD (status=1) and healthy subjects (status=0). The following observations can be made:

1. **Central Node - Status 1:**

- The node labeled "status=1" is central and connected to multiple rules, indicating that the majority of the association rules (from rule 1 to rule 10) are predictive of PD. The red color of these nodes highlights their significance in classifying PD cases.

2. **Prominent Features:**

- Features like mdvp.fo.hz (average vocal fundamental frequency), spread1, ppe, and mdvp.shimmer.db frequently appear across multiple rules, indicating their importance in predicting PD. These features have strong connections to multiple rules, further validating their critical role in classification.
- For example, mdvp.fo.hz and ppe appear in several rules such as rule 1, rule 2, rule 3, rule 4, and rule 7, emphasizing their strong association with Parkinson's.

3. **Interconnected Rules:**

- The arrows indicate how different rules share common features. For example, rule 1 and rule 2 both include mdvp.fo.hz, mdvp.shimmer.db, spread1, and ppe, showing strong overlap in feature sets. This suggests that there are specific combinations of vocal characteristics that are consistently associated with PD.

4. **Status 0 (Healthy):**

- Rule 11 is isolated and predicts the healthy status (status=0), showing much fewer associations compared to Parkinson's status. This reflects that fewer association rules are linked to healthy subjects, and distinguishing healthy individuals appears to be less complex in this dataset compared to detecting Parkinson's.

In summary, the network graph demonstrates a strong clustering of rules predicting PD with common vocal and dynamical complexity features, while the healthy status is represented by a singular, less connected rule. This visualization underscores the importance of specific vocal and dynamical features in the classification of PD.

Discussion

PD is a progressive neurodegenerative condition marked by the slow degeneration of dopaminergic neurons in the substantia nigra pars compacta, resulting in various motor and non-motor symptoms, including tremors, bradykinesia, stiffness, and cognitive deficits.Diagnosing PD is frequently difficult, as symptoms usually manifest only after considerable disease progression. This has propelled research towards creating techniques for earlier identification, including the examination of voice rhythms, which are frequently altered in the initial stages of PD. From a clinical perspective, early identification of PD is crucial as it enables physicians to initiate neuroprotective

DOI: 10.5455/medscience.2024.10.130

therapies and symptomatic treatments at a stage when they are most effective.

In this study, the successful application of associative classification for the diagnosis of PD using speech data has been demonstrated. The results show that the vocal characteristics of individuals with PD are strong indicators for early diagnosis. The classification algorithm achieved a high overall accuracy (0.928) and balanced accuracy (0.952), indicating its effectiveness in distinguishing between Parkinson's patients and healthy individuals. The high sensitivity rate (0.905) demonstrates the model's ability to reliably identify PD patients. These findings are consistent with earlier research that have effectively used machine learning algorithms to diagnose PD, showing speech analysis' promise as a non-invasive diagnostic tool [24-26].

The AC method employed in this study integrates association rule mining with classification techniques, allowing for the identification of significant relationships between vocal features and PD status. The rules generated, such as those involving fundamental frequency and shimmer measurements, underscore the importance of specific vocal characteristics in the classification process. This approach not only enhances interpretability but also facilitates the discovery of strong feature associations, which is crucial in medical diagnostics [27,28].

The findings clearly show the advantages of using associative classification in detecting PD based on vocal features. The use of "if-then" rules offers simpler and more interpretable results compared to traditional classification methods, making the model more practical for clinical use. For example, features such as fundamental frequency (MDVP), shimmer, spread1, and PPE were identified as key indicators of PD. The fact that these features can distinguish PD even before the onset of motor symptoms, as seen in vocal differences, makes this approach promising for early diagnosis. The ability to extract meaningful insights from complex datasets is a key advantage of this methodology, as it can lead to the identification of novel biomarkers for PD [29]. The classification rules obtained demonstrated a high degree of confidence (confidence=1) in accurately identifying individuals with PD. This indicates that the model provides very reliable predictions for those with the disease

However, the study also acknowledges certain limitations, particularly regarding the model's negative predictive value of 77.4%. This suggests that while the model excels in identifying positive cases, there is room for improvement in accurately predicting negative cases. Addressing this issue may involve refining the dataset through additional data collection and employing more sophisticated machine learning techniques, such as deep learning models, which have shown promise in similar contexts [30,31]. Future research could also explore hybrid approaches that combine multiple algorithms to enhance overall predictive performance [32].

Conclusion

In conclusion, the findings of this study contribute to the growing body of evidence supporting the use of machine learning for the early detection of PD through voice analysis. The high classification performance metrics achieved demonstrate the feasibility of this approach, while the insights gained from the associative classification method provide a valuable framework for future investigations into the vocal biomarkers of PD. Continued exploration in this domain holds the potential to significantly improve diagnostic accuracy and patient outcomes.

Conflict of Interests
The authors declare that there is no conflict of interest in the study.

Financial Disclosure
The authors declare that they have received no financial support for the study.

Ethical Approval
Since the dataset used in the study is open access, ethics committee approval is not required.

References

1. Deng Y, Liao Y, Huang P, et al. IRAK-M deficiency exacerbates dopaminergic neuronal damage in a mouse model of sub-acute Parkinson's disease. Neuroreport. 2023;34:463-70.
2. Armstrong MJ, Okun MS. Time for a new image of Parkinson disease. JAMA Neurol. 2020;77:1345-6.
3. Ullah MA. Unravelling Parkinson's disease prediction: an evaluation of feature selection techniques with a focus on PCA and KNN performance. Review of Computer Engineering Studies. 2023;10:20-7.
4. Mei J, Desrosiers C, Frasnelli J. Machine learning for the diagnosis of Parkinson's disease: a review of literature. Front Aging Neurosci. 2021;13;633752.
5. Real R, Martínez-Carrasco A, Lawton M, et al. Association between the LRP1B and APOE loci and the development of Parkinson's disease dementia. Brain. 2023;146:1873-87.
6. Schapira AH. Recent developments in biomarkers in Parkinson disease. Curr Opin Neurol. 2013;26:395-400.
7. Hely MA, Morris JG, Reid WG, Trafficante R. Sydney multicenter study of Parkinson's disease: Non-L-dopa–responsive problems dominate at 15 years. Mov Disord. 2005;20:190-9.
8. Mahajan A, Butala A, Okun MS, et al. Global variability in deep brain stimulation practices for Parkinson's disease. Front Hum Neurosci. 2021;15:667035.
9. Heim B, Krismer F, De Marzi R, Seppi K. Magnetic resonance imaging for the diagnosis of Parkinson's disease. J Neural Transm (Vienna). 2017;124:915-64.
10. Braak H, Ghebremedhin E, Rüb U, et al. Stages in the development of Parkinson's disease-related pathology. Cell Tissue Res. 2004;318:121-34.
11. Solana-Lavalle G, Rosas-Romero R. Analysis of voice as an assisting tool for detection of Parkinson's disease and its subsequent clinical interpretation. Biomedical Signal Processing and Control. 2021;66:102415.
12. Sakar BE, Isenkul ME, Sakar CO, et al. Collection and analysis of a Parkinson speech dataset with multiple types of sound recordings. IEEE J Biomed Health Inform. 2013;17:828-34.

DOI: 10.5455/medscience.2024.10.130

13. Lahmiri S, Shmuel A. Detection of Parkinson's disease based on voice patterns ranking and optimized support vector machine. Biomedical Signal Processing and Control. 2019;49:427-33.

14. Rayan Z, Alfonse M, Salem A-BM. Machine learning approaches in smart health. Procedia Computer Science. 2019;154:361-8.

15. Wan KR, Maszczyk T, See AAQ, et al. A review on microelectrode recording selection of features for machine learning in deep brain stimulation surgery for Parkinson's disease. Clin Neurophysiol. 2019;130:145-54.

16. Venturini L, Baralis E, Garza P. Scaling associative classification for very large datasets. Journal of Big Data. 2017;4:44.

17. Hadi We, Aburub F, Alhawari S. A new fast associative classification algorithm for detecting phishing websites. Applied Soft Computing. 2016;48:729-34.

18. Little M, Mcsharry P, Roberts S, et al. Exploiting nonlinear recurrence and fractal scaling properties for voice disorder detection. Biomed Eng Online. 2007;6:23.

19. Padillo FJ, Luna JM, Ventura S. Evaluating associative classification algorithms for big data. Big Data Analytics. 2019;4:2.

20. Thabtah F. A review of associative classification mining. The Knowledge Engineering Review. 2007;22:37-65.

21. Alataş B, Akin E. Rough particle swarm optimization and its applications in data mining. Soft Comput. 2008;12:1205-18.

22. Ayyagari MR. Integrating association rules with decision trees in object-relational databases. International Journal of Computer Trends and Technology. 2019;67:102-8.

23. Gonzalez-Abril L, Velasco F, Ortega JA, Cuberos FJ. A new approach to qualitative learning in time series. Expert Systems with Applications. 2009;36:9924-7.

24. Shahbakhi M, Far DT, Tahami E. Speech analysis for diagnosis of Parkinson's disease using genetic algorithm and support vector machine. Journal of Biomedical Science and Engineering. 2014;7:147-56.

25. Benba A, Jilbab A, Hammouch A, Sekkate S. Voiceprints analysis using MFCC and SVM for detecting patients with Parkinson's disease. 2015;3:300-4.

26. Suppa A, Costantini G, Asci F, et al. Voice in Parkinson's disease: a machine learning study. Front Neurol. 2022;13:831428.

27. Abiyev RH, Abizade S. Diagnosing Parkinson's diseases using fuzzy neural system. Computational and Mathematical Methods in Medicine. 2016;2016:1267919.

28. Gündüz H. Deep learning-based Parkinson's disease classification using vocal feature sets. IEEE Access. 2019;7:115540-51.

29. Abhishek MS, Chethan CR, Aditya CR et al. Diagnosis of Parkinson's disorder through speech data using machine learning algorithms. International Journal of Innovative Technology and Exploring Engineering. 2020;9:69-72.

30. Alzubaidi M, Shah U, Zubaydi HD, et al. The role of neural network for the detection of Parkinson's disease: a scoping review. Healthcare. 2021;9:740.

31. Chintalapudi N, Battineni G, Hossain MA, Amenta F. Cascaded deep learning frameworks in contribution to the detection of Parkinson's disease. Bioengineering. 2022;9:116.

32. Abba S, Isiyaku S, Boukari S, Lee J-A. Performance analysis and evaluation of supervised machine learning algorithms for the detection of Parkinson's disease. Preprints. 11 December 2023. doi: 10.20944/preprints202312.0687.v1