The Indonesian Journal of Computer Science
www.ijcs.net
Volume 14, Issue 2, April 2025
https://doi.org/10.33022/ijcs.v14i2.4837

Machine Learning and Explainable AI for Parkinson's Disease Prediction: A Systematic Review

Belinda Ndlovu[1], Kudakwashe Maguraushe[2] , Otis Mabikwa[3]
belinda.ndlovu@nust.ac.zw[1], maguraushe.kuda@mut.ac.za[2], n02133973d@nust.students.ac.zw[3]
[1,3] National University of Science and Technology, Ascot AC 939, Bulawayo
[2] Mangosuthu University of Technology, 511 Griffiths Mxenge Hwy, Umlazi, Durban, 4031

Article Information

Received : 8 Apr 2025
Revised : 22 Apr 2025
Accepted : 30 Apr 2025

Keywords

Parkinson's disease, machine learning, predictive, artificial intelligence, Explainable Artificial Intelligence.

Abstract

Parkinson's disease is a movement disorder within the nervous system that impacts millions of people across the world. The standard diagnostic methods usually miss early subtle signs of the disease, which has driven research into Machine Learning (ML) and Explainable Artificial Intelligence (XAI) to develop better predictive models. Following PRISMA guidelines we analyzed 13 studies found in IEEE Xplore, PubMed and ACM concerning different ML methodologies for Parkinson's disease prediction. Deep learning models using vocal and motor data achieve good accuracy but require more clinical trust and transparency due to their opaque "black-box" nature. SHAP and LIME act as XAI solutions that address transparency issues in model predictions by delivering understandable information regarding model outputs to all users. Current solutions show progress. However, there are multiple complications including limited and unbalanced datasets alongside accuracy-explainability trade-offs which underline the need for extensive datasets, multidisciplinary teamwork and practical validation.

A. Introduction

Parkinson's Disease (PD) is a progressive disease of the nervous system which produces symptoms such as tremors, stiffness and walking problems which leads to reduced quality of life, mainly affecting middle-aged and elderly people [1]. The exact cause of PD remains unknown although data shows over 10 million people across the globe suffer from it, with those above 50 years old being particularly affected [2], [3]. Traditional PD diagnosis methods have some limitations, they depend on subjective clinical evaluations and struggle to detect the disease early [1], [4]. Traditional diagnosis includes evaluating medical history and performing physical examinations while analyzing clinical symptoms like tremors, rigidity, bradykinesia and movement slowness [1], [4], [5]. The human eye generally cannot detect these symptoms because they are typically too subtle which leads to misdiagnosis [1], [6]. The existing traditional methods fail to provide clear-cut blood biomarker tests and neuroimaging scans for early detection of PD [7], [8]. Furthermore, PD has symptoms that overlap with other neurodegenerative diseases such as essential tremor and Alzheimer's [6]. Resultantly, this also leads to potential misdiagnosis [1]. This results in PD being diagnosed at much later stages when neuronal damage has been done, narrowing the scope for available treatments [4], [8].

Machine Learning (ML) is a subset of Artificial Intelligence (AI) that focuses on the use and development of computer systems that can learn, adapt, recognize patterns and make predictions without being explicitly programmed for every task [9], [10], [11]. This presents the opportunity for earlier and more accurate detection of PD. ML models have achieved high accuracies when predicting chronic diseases such as diabetes [12][13] , Human Immunodeficiency Virus (HIV) [14] and PD [9], with some exceeding 90% accuracy.

ML models predict and diagnose PD by feature selection and extraction which helps identify patterns and potential biomarkers, using data from various sources such as wearable sensors and smartphones [9]. Despite the promise ML has in diagnosing PD, challenges remain in integrating these tools into clinical settings such as lack of explainability, limited data and interoperability issues, however, opportunities lie in developing more transparent models and accessible self-assessment tools [15], [16], [17].

Explainable Artificial Intelligence presents an opportunity to counter the challenges and limitations of ML [18]. Explainable Artificial Intelligence refers to the techniques and methods in AI that have the objective to explain the decision-making process behind AI models, making them more transparent, interpretable, and easy to comprehend to humans thus increasing medical personnels' trust [19], [20]. This is crucial for trust, accountability, and ethical considerations, particularly in healthcare [21], [13], [22]. The intersection and combination of ML and XAI in PD diagnosis presents more accurate, more objective, and sensitive prediction while enabling early detection mechanisms through data analysis [4].

A comprehensive review of ML algorithms in PD diagnosis was done by [5] and [2]. The literature reviewed ML and Deep Learning (DL) applications for improving the diagnosis of PD from the data modalities of speech, handwriting, gait, and neuroimaging. The focus was on classifying PD and healthy controls using AI, ML, and DL applications, whilst exploring the use of voice data and handwritten

patterns. The articles provided evidence to prove that ML and DL can assist in carrying out diagnosis for PD from various data types. Key challenges in both studies were found to be data limitations, including small, imbalanced datasets, the need for clinical validation, and the interpretability of models developed using ML. Nevertheless, the areas of opportunity for the future included diagnosis by real-time customized devices, Big Data Analytics applications, and either developing or improving methods for dealing with missing and multimodal data.

Automated PD diagnostics using DL models were explored in [1] and [4]. Their work covered different types of data modalities, including brain analysis (MRI, SPECT, EEG, and PET) and motor indicators including speech, handwriting, gait, and EMG. They also proved that DL could enhance PD disease diagnosis and facilitate informed decision-making for clinicians. The DL models demonstrated high accuracy in diagnosing PD on a variety of modalities. However, the limitations of these DL models were their "black box" nature, the need for clinical trials and poor interoperability in the use of different datasets and the difficulty in obtaining regulatory approval.

This systematic literature review (SLR) will examine the most commonly used algorithms for PD prediction alongside XAI, addressing both performance and transparency in predicting PD. By critically evaluating how these ML models handle clinical challenges such as data quality, interpretability, and bias, it not only synthesizes current research but also identifies gaps and opportunities for improving diagnostic reliability and trust.

Research Questions

1. Which ML algorithms are most commonly used for predicting Parkinson's disease?
2. How does XAI contribute to the interpretability and trustworthiness of predictions in Parkinson's disease?
3. What are the main challenges in current research on ML and XAI for Parkinson's disease prediction?
4. How do ethical and privacy concerns impact the clinical application of ML and XAI in predicting Parkinson's disease, and what strategies can be implemented to mitigate these issues?

The rest of this paper is structured as follows: Section 2 outlines the methodology, Section 3 the results and Section 4 the discussion that presents the detailed analysis of the research results and identifies research gaps for future study.

B. Methodology

The study adopted Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines to guarantee the quality and rigor of the study. These guidelines included the identification, screening, and eligibility criteria for the literature review, as well as the structured analysis of the findings [23].

Search strategy

A widespread search was conducted on 24 November 2024 on the following three databases: IEEE Xplore, PubMed, and ACM. A combination of keywords and their synonyms was used to formulate a search strategy which was modified to suit each database syntax as follows: ("Parkinson's disease" OR "Parkinson disease" OR "PD") AND ("machine learning" OR "ML" OR "artificial intelligence" OR "AI") AND ("Explainable AI" OR "Explainable Artificial Intelligence" OR "XAI") AND ("prediction" OR "diagnosis" OR "detection") AND ("algorithms" OR "models" OR "methods").

Inclusion and exclusion criteria

Table 1. Inclusion and exclusion criteria.

Criteria	Inclusion	Exclusion
Time frame	2020 to 2024. To provide the most recent scholarly research in the field.	Studies published farther than five years ago.
Language	Only English.	Any other language that is not English.
Type of paper	Conference proceedings and Journal articles	Grey literature, book chapters, commentary pieces and editorials
Publication status	Final publications.	Studies and Abstracts in draft form or that have not been peer reviewed.
Research area	The application of application of ML and XAI in predicting Parkinson's disease.	Studies that were outside the span of this articles' research area.
Keywords	Machine learning, predicting, Parkinson's disease, explainable artificial intelligence	Studies not explicitly focusing on these topics in the context of predicting Parkinson's disease.

Eligibility and screening

The initial search yielded a total of 14 articles from IEEE Xplore, 13 articles from PubMed, and 1 article from ACM. Thus, a total of 28 papers met the Eligibility standards, based on whether they were peer-reviewed, published in journals or conferences. After this initial search, the 28 studies underwent title and abstract screening. During this stage, n=9 studies were excluded because they did not focus on ML predicting PD and were not able to answer the study's formulated research questions, leaving n=19 studies for full-text review.

Included

The full-text assessment led to the further exclusion of n=6 studies that did not meet the eligibility criteria, primarily due to a lack of focus on ML applications in Parkinson's disease and XAI. Ultimately, a total of n=13 studies were included in the final synthesis. The flow diagram for this study using PRISMA is depicted in Figure 1.

C. Results

The delimitation process is shown in the following PRISMA flowchart by [23].

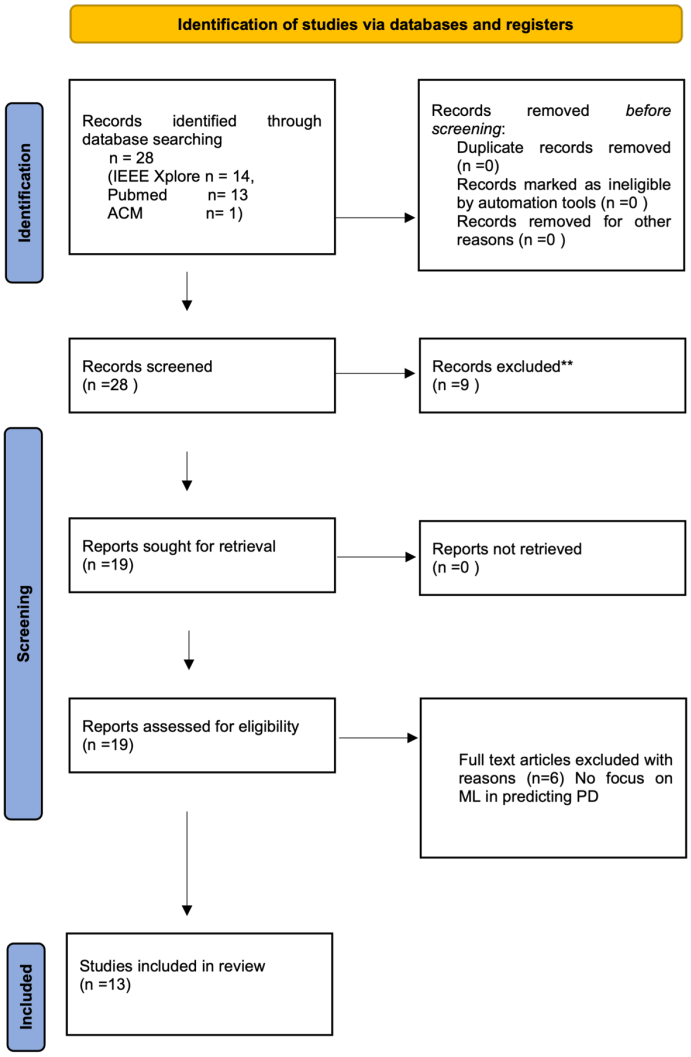

Figure 1. The search process with inclusion and exclusion criteria

Data Reporting

A seven-column table was created to explore research questions and summarize article findings. This study analyzed 13 papers. Table 2 presents the comprehensive list of publications and their corresponding factors. Table 2 also categorizes the 13 articles based on the author's name(s), the country of research, the most reliable algorithm, the study's opportunities, its challenges or limitations, the algorithms or techniques employed, and the summary of the findings.

Table 2. Characteristics of the included articles.

Author	Origin	Accuracy	Opportunities	Challenges	Algorithms	Main Findings
[24]	USA	LightGradient-Boosting Machine (LightGBM 0.96%)	▪ Potential for Personalized Medicine ▪ Early diagnosis ▪ Accurate Diagnosis ▪ Subtype Prediction ▪ Progression Analysis ▪ Exploration of New Markers	▪ Limited Dataset Size ▪ Data Imbalance ▪ Generalizability ▪ Real-World Application	▪ LightGBM ▪ Extreme Gradient Boosting (XGBoost) ▪ AdaBoost ▪ Bagging (Bootstrap Aggregating ▪ Support Vector Machine	▪ Exceptional Sensitivity and Specificity of LightGBM ▪ Vocal Biomarkers as Promising Indicators ▪ ML is a valuable tool for analyzing complex patterns
[25]	USA	Generalized Forest (gcF) 0.94%	▪ non-invasive ▪ Cost-effective Diagnosis ▪ Early Detection ▪ Monitoring ▪ Personalized treatment strategies ▪ Development of computational tools ▪ Integration of multiple biomarkers	▪ Data Dependency ▪ Computational Resources ▪ Training Time ▪ Generalizability ▪ Interpretability ▪ Focus on Speech Features	▪ Generalized Forest (gcF), Logistic Regression, ▪ Support Vector Machines (SVM), ▪ XGBoost, ▪ LightGBM, ▪ CatBoost ▪ neural networks ▪ Gradient Boosting Machines (GBM)	▪ Influence of Feature Selection (PPE and RPDE exhibit a strong correlation with the target variable PD status) ▪ Impact of Hyperparameter Tuning for the good ▪ Performance Variations in Neural Networks ▪ Potential of Raw Voice Data

[26]	Korea	Neural Network 0.77%	▪ Early Detection and Intervention ▪ Improving Diagnostic Accuracy ▪ Personalizing Treatment Strategies ▪ Facilitating Large-Scale Screening ▪ Developing Cost-Effective Solutions ▪ Supporting Clinical Decision Making	▪ Generalizability to Other Populations ▪ Potential Selection Bias ▪ Recall Bias in Self-Reported Data	▪ Logistic Regression ▪ Random Forest (RF) ▪ Neural Network ▪ Gradient Boosting Machines (GBM) ▪ Decision Tree ▪ Naïve Bayes ▪ XGBoost	▪ BMI is a Strong Predictor, ▪ Lifestyle Factors Play a Role ▪ Sex Differences in Predictive Factors ▪ Lifestyle factors, including smoking and alcohol consumption, were more strongly associated with PD risk in men ▪ Cost-Effective Screening Using Existing Data
[27]	India	Gradient Boosting 0.96%	▪ Development of Larger datasets ▪ More Diverse Datasets ▪ Enhancing Model Interpretability ▪ Enhancing Trustworthiness ▪ Integration of Multiple Data Modalities ▪ Longitudinal Studies for Validating Biomarker Significance ▪ Focus on Translating Research to Clinical Practice ▪ Exploring New AI Techniques and Applications ▪ Developing Patient-Centric AI Solutions	▪ Reliance on a Specific Toolkit for MRI Processing ▪ Focus on a Single ML Mode ▪ Limited Scope of Data Modalities ▪ Lack of a User-Friendly Interface ▪ Need for Further Clinical Validation	▪ K-Nearest Neighbours (KNN) ▪ Extra Trees (ET) ▪ RF ▪ CatBoost. ▪ Decision Tree (DT)Gradient Boosting (GB)	▪ Gradient Boosting as the Top-Performing Mode ▪ Identification of Significant Radiomics Features ▪ Enhancing Transparency and Trust with XAI, Addressing Dataset Imbalance by SMOTE

[28]	India	VGG19-INC Model 0.98%	▪ Construction of a precise deep learning algorithm to detect PD early ▪ Non-invasive ▪ Cost-Effective ▪ utilizing LIME to solve the classification problem ▪ Improving diagnostic accuracy	▪ Limited Dataset Size ▪ Data Variability ▪ imbalance of data ▪ deep learning models are difficult to explain	▪ convolutional neural networks (CNNs) ▪ AlexNet ▪ VGG19 Net ▪ ResNet-50 ▪ DenseNet-201 ▪ SqueezeNet 1_1, VGG19-INC ▪ LIME	▪ Differential Learning Rates Improve Model Performance ▪ Spiral, and Wave Drawings as Potential Biomarkers ▪ The VGG19-INC model achieved an accuracy of 98.45% ▪ ResNet-50 accuracy of 98.3% ▪ LIME enhanced the transparency and trustworthiness of the model
[29]	India	Logistic Regression 0.85%	▪ Identification of Risk Factors and Preventative Measures ▪ Non-invasive and Cost-effective Biomarkers	▪ Data Quality and Availability ▪ Generalisability and Validation	▪ RF ▪ SVM ▪ Logistic Regression	▪ Early Detection is Crucial for Effective Treatment and Management
[30]	USA	Multimodal Model (AdaBoostClassifier)0.90%	▪ Drug Discovery ▪ Drug development ▪ Personalized Treatment Strategies ▪ Understanding Disease Mechanisms ▪ Expanding to Diverse Population ▪ Incorporating Additional Predictors ▪ Open Science and Collaboration	▪ Data Variability and Generalisability ▪ Lack of diversity in available sample series ▪ Lack of an optimal dataset to validate the findings ▪ Focus on Early-stage PD	▪ Stochastic Gradient Descent (SGDClassifier) ▪ K-Nearest Neighbours ▪ Logistic Regression ▪ Adaptive Boosting ▪ Support Vector Machines (SVC) ▪ Multi-layer Perceptron	▪ Multimodal Approach Outperforms Single Modality ▪ UPSIT, and PRS as Key Predictors ▪ AdaBoostClassifier is the Most Accurate Algorithm ▪ Tuned Model Achieves High Accuracy ▪ Gene Expression Network Communities Offer Insights ▪ Open Science Framework Promotes Reproducibility and Collaboration

					Neural Networks ▪ Linear Discriminant Analysis ▪ Gradient Boosting	
[31]	India	XGBoost with Recursive Feature Elimination (RFE), 0.96%	▪ Early diagnosis ▪ Accurate Diagnosis ▪ Personalized Treatment Planning ▪ Non-Invasive and Accessible Screening ▪ Further Model Refinement ▪ Further Model Validation ▪ Multimodal Integration, ▪ Real-Time Applications	▪ Limited Dataset Size ▪ Potential Bias in Synthetic Data ▪ Applicability in Real-Time Settings ▪ Ethical Considerations ▪ Clinical Validation	▪ RFE with XGBoost, ▪ SVMSMOTE ▪ SHAP ▪ KNN, ▪ RF, ▪ LR ▪ DT ▪ MLP ▪ Gaussian Naive Bayes	▪ XGBoost Classifier with Recursive Feature Elimination (RFE) achieved the highest accuracy (96.61%) for PD prediction, ▪ Pitch Period Entropy (PPE) was identified as the most important feature for PD prediction ▪ Addressing class imbalance through SVMSMOTE significantly improved model performance, ▪ The integration of SHAP (SHapley Additive exPlanations) gave valuable insights into the model's decision-making process
[32]	Italy	DenseNet + EN 0.96%	▪ Early detection ▪ Pre-symptomatic Detection ▪ Improved Diagnostic Accuracy ▪ Improved objectivity ▪ Personalized Treatment ▪ Monitoring	▪ Data Dependency ▪ Potential Overfitting ▪ Challenges in Model Interpretability ▪ Limited Dataset Diversity	▪ DenseNet 3D ▪ Vision Transformer 3D (ViT) ▪ ResNet 3D ▪ RF ▪ Excitation Network (EN) ▪ SVM ▪ XGBoost	▪ DenseNet with EN Emerged as the Best-Performing Model (F1-score of 96.5%) ▪ XAI Revealed Clinically Relevant Insights ▪ Lateral Ventricle Enlargement as a Potential Prodromal Indicator

			▪ Discovery of Novel Biomarkers ▪ Discovery of disease Mechanisms ▪ Development of Accessible and Cost-Effective Diagnostic Tools	▪ Focus on Specific Modalities ▪ Need for Longitudinal Studies ▪ Translating Research into Clinical Practice		▪ Multimodal Learning Outperforms Unimodal Approaches ▪ Bradykinesia as a Key Discriminating Feature
[33]	India	SVM 0.94%	▪ High accuracy in prediction ▪ Personalized treatment ▪ Non-invasive diagnosis ▪ Early detection ▪ Large-scale Screening and Accessibility ▪ Supporting Clinical Decision-Making	▪ Clinical validation ▪ Data variability ▪ Data Generalisability and Bias ▪ Reliance on Self-Reported Data ▪ Lack of Detailed Clinical Data	▪ Naïve Bayes ▪ SVMs ▪ Networks (ANNs) ▪ RF ▪ k-Nearest Neighbours (k-NN) ▪ Artificial Neural	▪ ML demonstrates significant potential for PD prediction ▪ Voice and speech data emerge as promising sources for non-invasive diagnosis. ▪ Accuracy levels vary across studies and algorithms
[34]	China	PD-ResNet model 95.51%	▪ Developing Real-Time Personalised Devices ▪ Developing a Comprehensive ML Model ▪ Expanding Wearable Sensor Capabilities ▪ Implementing Cloud- and ML-Based Frameworks ▪ Long-Term Monitoring in Real-Life Conditions	▪ Manifold Modelling ▪ Model Interpretation ▪ Lack of Clinician Trust ▪ Addressing Imbalanced Datasets ▪ Limited Sensor Capabilities ▪ Differential Diagnosis	▪ PD-ResNet ▪ SVM ▪ GoogLeNet ▪ RF ▪ XGboost	▪ SMOTE addressed imbalanced datasets ▪ Polynomial elevated dimensions technique ▪ PD-ResNet outperformed other algorithms

			▪ Addressing Freezing of Gait (FOG) ▪ Differential Diagnosis with Other Neurological Disorders			
[35]	India	XGBoost 0.96%	▪ Potential for Novel Biomarker Discover ▪ Non-Invasive Diagnosis ▪ Real-world Application ▪ Real world Integration ▪ Personalized Treatment Strategies ▪ Expanding Datasets ▪ expanding Features ▪ Developing More Accurate Diagnostic Tools	▪ Generalizability to Real-World Settings ▪ Interpretability of ML Models	▪ Logistic Regression ▪ XGBoost	▪ Audio Signals as a Potential Diagnostic Tool ▪ Twelve Key Features Identified ▪ XGBoost outperformed Logistic Regression
[36]	Saudi Arabia	ensemble deep learning network (DEEP_EN) 0.96.%	▪ Objective and Quantifiable Assessment ▪ Improved Research ▪ Drug Development ▪ Early detection of PD	▪ Limited Data Availability ▪ Data Complexity and Variability ▪ interpretability and Explainability of Models	▪ Feed-Forward Neural Networks ▪ Classification Trees ▪ RF ▪ BOOST_TREE:	▪ importance of Specific Premotor particularly the striatal binding ratios ▪ Superior Performance of Deep Learning ▪ Boosting algorithms demonstrated comparable performance to deep learning,

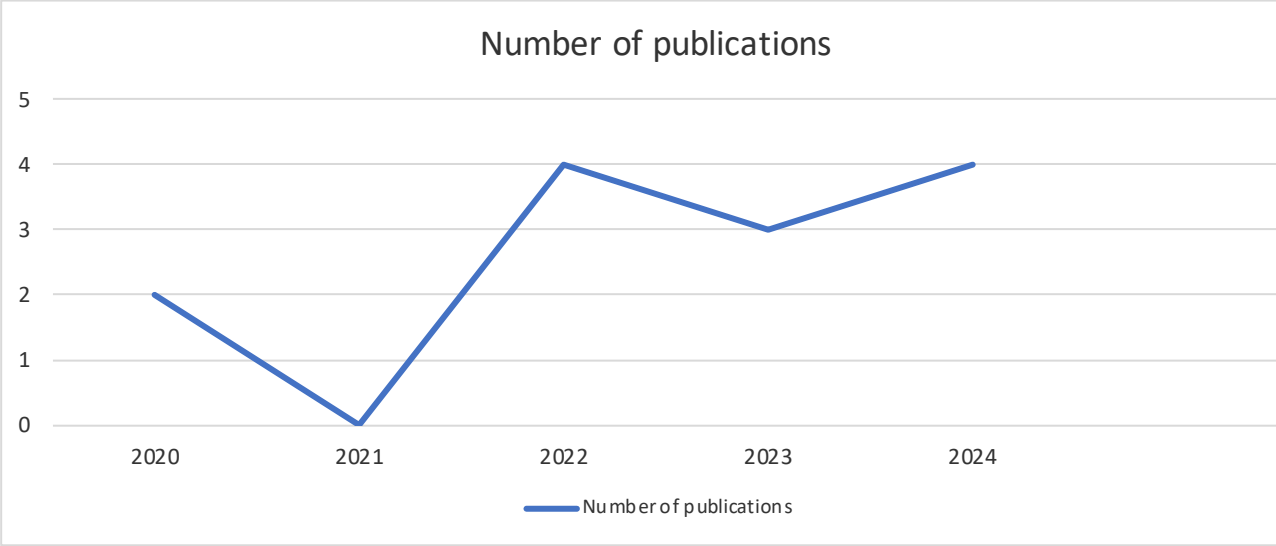

Figure 2. Number of Publications

Figure 2 shows the publication trend of the papers reviewed in this SLR. Despite a slight decrease from 2020 to 2021, the graph shows a growing interest in ML and XAI from 2021 in PD prediction. The fluctuations in the subsequent years show that ML and XAI in PD prediction have not yet matured in healthcare and are likely fueled by advancements in technology, improved datasets, and a rising interest in AI-driven medical solutions.

Study characteristics and origin

Table 3 shows an analysis of the 13 selected studies in this SLR. It reveals that Asia is the leading continent, with its research articles making up 69% of the papers presented in this SLR, followed by America (23%), Europe (8%), and then Africa and Oceania with limited involvement in this technology.

Table 3. Study characteristics and origin

Continent	***Number of papers***	***Percentage***
America	3	23%
Asia	9	69%
Europe	1	8%
Africa	0	0%
Oceania	0	0%

Opportunities

Table 4. Opportunities

Opportunity	Source	Description	Potential Impact
Early Detection	[27], [24], [25], [28], [36], [34], [29], [30], [26], [31]	ML models can detect subtle patterns in data that suggest the early onset of PD.	Enables timely interventions, slowing disease progression.
Personalized Medicine	[30], [27], [25], [32], [30], [29]	Tailored predictions based on individual patient profiles and biomarkers.	Improves treatment plans and outcomes for patients.
Multimodal Data Integration	[25], [32], [30], [28], [31], [29], [32], [26], [36]	Combines diverse data sources like vocal, motor, imaging, and genetic data.	Enhances prediction accuracy and provides a holistic view of disease progression.
Explainability (XAI)	[28], [27], [32], [31], [30]	XAI methods offer insights into how predictions are made by ML models.	Builds trust with clinicians and patients, facilitating clinical adoption.
Automation in Screening	[27], [25], [33], [29], [32], [36], [28]	ML-based tools automate initial screening processes using non-invasive data like voice analysis.	Reduces the burden on healthcare systems and improves accessibility.
Improved Accuracy	[27], [24], [25], [30], [34], [31], [30], [29], [26], [34], [35]	Advanced algorithms (e.g., deep learning) provide higher predictive accuracy.	Reduces diagnostic errors and ensures reliable predictions.
Resource Allocation	[30], [28], [30], [36], [34], [32]	Predictive insights help allocate healthcare resources more efficiently.	Improves cost effectiveness and prioritization of care for high-risk patients.
Advancing Research	[24], [25], [30], [28], [36], [27], [30], [31], [29], [34],	ML can uncover new biomarkers and patterns related to PD progression.	Drives innovation and further understanding of the disease.

Table 4 highlights the key opportunities in using ML for PD prediction. These insights show how ML and XAI could enhance diagnostic accuracy, improve patient outcomes and facilitate trust through XAI. Improved Accuracy is at the top of the list with 11 citations, followed by advancing research and early detection cited by 10 authors each. Multimodal data Integration was cited by 9 authors followed by Automation in screening that was cited by 7 authors. Personalized medicine and resource allocation were cited by 6 authors each and Explainability was cited by 5 authors. This spread suggests that while strengthening overall research and clinical practices is a major focus, areas like personalized medicine and explainability might need more attention. These advancements underscore the potential ML has to revolutionize PD diagnosis and management while supporting clinicians in delivering more effective, timely and tailored care.

Challenges

Table 5 outlines the challenges associated with using ML for PD prediction. These insights show the need for standardized practices, real-world validation and ethical data handling.

Table 5. Challenges

Challenge	Author	Description	Impact
Data Availability	[30], [27], [36], [32], [28]	Limited access to large, high-quality, and diverse datasets.	Reduces model generalizability and accuracy.
Class Imbalance	[27], [24], [31], [34]	Datasets often have fewer samples of PD-positive cases compared to healthy controls.	Leads to biased models and poor performance on minority classes.
Heterogeneous Data	[32], [30], [26], [36]	Variability in data sources (e.g., imaging, voice, genetic data) makes integration challenging.	Complicates feature extraction and model development.
Overfitting	[30], [31], [29], [27], [25], [28], [34], [32], [24], [36]	Models may perform well on training data but poorly on unseen data.	Limits real-world applicability.
Explainability	[27], [28], [32], [31]	Many high-performing models (e.g., deep learning) lack transparency in their decision-making.	Reduces clinician trust and hinders clinical integration.
Ethical and Privacy Concerns	[32], [30]	Handling sensitive patient data raises privacy and consent issues.	Restricts data sharing and collaboration.
Data Preprocessing	[29], [32], [28], [30], [31], [36], [34], [24],	Complex preprocessing is required for noisy and incomplete datasets.	Increases computational effort and processing time.
Lack of Standardization	[27]	No universal standards for feature selection, evaluation metrics, or data formats.	Makes comparisons between studies difficult.
Real-World Validation	[27], [32], [28]	Limited real-world testing and validation of predictive models.	Raises concerns about the clinical utility of the models.
Computational Complexity	[27], [36],	Advanced models like deep learning require significant computational resources.	Limits accessibility for smaller research teams or resource-limited settings.

Overfitting is the most highlighted challenge cited by 10 authors, indicating that models frequently have difficulty generalizing new data. Challenges in data preprocessing come in second with 8 citations while data availability has 5 citations. Heterogenous data, class imbalance and explainability have 4 citations each. The lack of real-world validation has 3 citations, computational complexity 2 citations and lack of standardization with 1 citation.

Algorithm

Table 6 highlights the barriers to clinical adoption. Addressing these challenges is essential for successfully integrating ML and XAI into clinical workflows.

Table 6. Algorithms used

Algorithm	Author
Data Preprocessing Techniques:	
Extremely Randomised Trees Classifier Algorithm (extraTrees)	[30]
Synthetic Minority Over-sampling Technique (SMOTE)	[31]
Recursive Feature Elimination (RFE)	[31]
Principal Component Analysis (PCA)	[31]
Correlation-based Feature Selection (CFS)	[33]
Feature Encoding with Evolutionary Wavelet Neural Networks (EWNNs)	[31]
Mel-Frequency Cepstral Coefficients (MFCCs) with Deep Neural Networks (DNN)	[31]
Polynomial Elevated Dimensions Technique	[34]
Machine Learning Algorithms	
Neural Networks:	
Artificial Neural Networks (ANN): Including Multilayer Perceptron (MLP) and Radial Basis Function (RBF) Networks	[33], [28]
Deep Neural Networks (DNN)	[31]
Convolutional Neural Networks (CNN)	[31], [28]
Recurrent Neural Networks (RNN)	[25]
Long Short-Term Memory (LSTM)	[25], [36]
Multilayer Perceptron (MLP)	[30], [31]
Support Vector Machines (SVM)	[24], [25], [29], [30], [31], [32], [33], [34]
Bayesian Models:	
Bayesian Belief Network (BBN)	[33]
Naïve Bayes	[33], [29]
Tree-Based Models:	
Decision Trees	[29], [26]
Random Forest	[26], [27], [29], [31], [32], [33], [34], [36]
Classification and Regression Trees (CART)	[33]
Extremely Randomized Trees	[30]
Boosting Methods:	
AdaBoost	[30], [24]
Gradient Boosting Machine (GBM)	[25], [26], [30]
eXtreme Gradient Boosting (XGBoost)	[24], [25], [26], [31], [32], [34], [35]
LightGBM	[24] [25]
Linear Models:	
Logistic Regression	[25], [26], [29], [30], [31], [34], [35]
Linear Discriminant Analysis (LDA)	[30], [35]
Quadratic Discriminant Analysis	[30]
Other Algorithms:	

K-Nearest Neighbors (KNN)	[27], [30], [33]
Adaptive Neuro-Fuzzy Classifier (ANFC).	[33]
Rotation Forest (RF)	[33]
Bagging Classifier	[30]
Discriminative Deep Forest	[25]
Fuzzy k-Nearest Neighbor (FKNN).	[33]
Extreme Learning Machine (ELM)	[24]
Ensemble Averaging	[29]
Explainable AI (XAI) Techniques:	
SHapley Additive exPlanations (SHAP)	[31], [30]
Local Interpretable Model-Agnostic Explanations (LIME)	[28]
Integrated Gradients (IG)	[32]
Attention Heatmaps	[32]
Feature Importance	[32]

Table 6 presents an overview of the machine learning algorithms employed in various studies for Parkinson's disease prediction, along with the respective authors who implemented them.

D. Discussion

This section examines the extant body of literature and answers the research question posed in this study.

1. The most commonly used algorithms for predicting Parkinson's disease

Multiple ML algorithms were found in this research and they demonstrated effectiveness in predicting PD. It was found that the effectiveness of an algorithm can depend on the type of data used, such as speech, gait, handwriting, MRI, or a combination of those [37], [38]. Therefore, no one algorithm performs perfectly across all data types or datasets. Thus, some studies explore the use of XAI techniques to see how features affect predictions thereby improving model transparency [39]. This means the choice of algorithm depends on the data available and the desired balance between accuracy, sensitivity and specificity.

Data preprocessing and feature selection algorithms

Data preprocessing and feature selection are key in developing effective ML models for PD prediction [31]. Techniques such as the Synthetic Minority Over-sampling Technique (SMOTE) help balance datasets, so that minority classes (often early-stage Parkinson's) are well represented [24]. For the analysis of voice recordings, one of the most important PD biomarkers, feature extraction methods such as Mel-Frequency Cepstral Coefficients (MFCCs) with Deep Neural Networks (DNN) are used for analysis [33], while dimensionality reduction methods such as Principal Component Analysis (PCA) and Correlation-based Feature Selection (CFS) are employed to eliminate noise and redundancy in high-dimensional datasets [33]. Furthermore, to rank and select the most relevant features in the data so the model is minimized and more accurate, algorithms such as Extremely Randomized Trees (extraTrees) and Recursive Feature Elimination (RFE) are used [30].

These approaches have their own strengths and weaknesses. For instance, techniques like SMOTE work on addressing class imbalance but such introduction of synthetic data can sometimes distort the underlying data distribution if not well regulated [31]. Methods like PCA that focus on dimensionality reduction and RFE, or CFS, that focus on feature selection, while successful in moderating overfitting and reducing the computational cost may remove subtle but clinically significant features [25]. Furthermore, enhanced approaches such as feature encoding using Evolutionary Wavelet Neural Networks (EWNNs) and Polynomial Elevated Dimensions Technique provide innovative methods to identify sophisticated patterns in data however, they need extensive computational resources and require meticulous parameter tuning [36]. Overall, these preprocessing and feature selection techniques allow numerous opportunities to enhance the accuracy and interpretability of predictive models in PD prediction, yet they demand a cautionary balance with a focus on clinical significance for potential applications.

Deep learning (DL) algorithms

These algorithms are a subfield of ML methods that use Artificial Neural Networks (ANN) with representation learning [40]. They are used to analyze complex relationships within multimodal data, including neuroimaging, genetics, clinical and demographic information [32]. DL algorithms can analyze complex relationships within multimodal data, for example [36] designed a deep learning model to discriminate between normal individuals and PD patients based on premotor features such as Rapid Eye Movement (REM) sleep Behavior Disorder (RBD) and olfactory loss, achieving a 96.45% accuracy.

These approaches still present many challenges. One of the main challenges is interpretation, considering that deep learning models often function as "black boxes" that offer little in terms of understanding or explaining their predictions [41]. The lack of interpretability makes it difficult for healthcare professionals to trust and apply the model's findings in practice [42]. An even greater challenge lies in the need for high-quality balanced datasets, as data imbalances and individual variations in the symptoms of PD may affect the performance of the model [5]. Although data augmentation methodologies increase data samples, improper augmentation may considerably reduce the performance of the network [40]. Despite these challenges, DL algorithms offer great opportunities for improving early PD detection, enabling timely interventions and better management of the disease.

Traditional ML algorithms

Traditional ML algorithms can be utilized in the analysis on highly complex datasets to extract important biomarkers and patterns necessary for PD prediction [40]. Algorithms such as Logistic Regression, SVM, Random Forests, and Decision Trees are then applied on diverse data modalities such as: clinical, genetic, and imaging, then speech and gait features [30]. The aforementioned classic algorithms are used in developing predictive models that can discriminate against healthy individuals from patients suffering from PD [26]. Furthermore, these algorithms can also handle processing of large complex datasets, finding patterns, and relationships among different variables [29] such that they give very useful knowledge on PD

classification and diagnosis with a lower cost [30]. Despite the opportunities that these algorithms present, they also have challenges such as feature selection [43]. While identifying the most relevant features is critical for accurate prediction, having more features can lead to a downturn in model precision due to irrelevant or correlated feature subsets [31]. Some traditional ML models can lack interpretability, making it difficult to understand the reasoning behind their predictions [40].

Explainable Artificial Intelligence (XAI) algorithms

XAI techniques are used in ML to provide transparency and interpretability to complex ML models, helping understand why a model makes certain predictions, which is crucial in healthcare where decisions can be life-altering [32]. These techniques include SHapley Additive exPlanations (SHAP) and Local Interpretable Model-Agnostic Explanations (LIME), and they are applied to demystify "black box" models like deep neural networks [31]. In a study by [32], Integrated Gradients were used for ResNet (Residual Network) and DenseNet (Densely Connected Convolutional Network), while Attention Heatmaps were used for Vision Transformer (ViT) to pinpoint regions where networks focused most. The use of XAI presents multiple opportunities such as enhanced trust and explainability [28]. XAI, by explaining model decisions, increases trust among healthcare professionals, thereby making them more likely to adopt AI-driven diagnostic tools [31]. These techniques can reveal critical biomarkers that may be overlooked by traditional methods, potentially leading to more accurate and earlier diagnosis [28]. Through its capacity to analyse complex data such as brain regions or specific symptoms critical to prodromal PD pathophysiology, such as the right temporal and the left pre-frontal areas, XAI reveals clinical insights that guide further research and clinical assessments [30].

Despite its potential, XAI also faces challenges such as complexity. Implementing and interpreting XAI methods can be complex, requiring expertise in both machine learning and the specific domain of application [31]. While some XAI techniques are computationally intensive, particularly when applied to large datasets or complex models [36]. Furthermore, XAI techniques hinge on the quality and quantity of the training data but if utilised on biased or incomplete data, they can lead to misleading explanations and unreliable predictions [32]. There is also limited generalizability of explanations, meaning some XAI methods may be specific to individual predictions or local regions of the input space, making it difficult to generalise these explanations to the entire model or dataset, which would lead to more reliable, transparent, and clinically useful diagnostic tools [28].

2. Opportunities of ML in PD Prediction

There are multiple opportunities for advancing PD research and care using ML techniques [44]. These opportunities stem from the growing potential of these technologies to improve diagnosis, prognosis, and treatment strategies for PD [29]. Early and Accurate Diagnosis with pre-motor Symptom Detection is a possibility with ML as pre-motor features such as olfactory loss, rapid eye movement, sleep behavior disorder, and subtle changes in vocal patterns can be analyzed [45]. ML algorithms enable earlier diagnosis and enable early therapeutic interventions

beneficial in slowing the progression of degenerative disease [46]. An impressive accuracy of 96.45% was reported by [36] in their study that used a deep-learning model with pre-motor features as biomarkers to detect early PD. Hence, ML models provide objectivity and quantification with higher repeatability when compared to traditional clinical examinations, which are largely subjective [24]. This can improve diagnostic accuracy and consistency, particularly in cases where symptoms are too subtle for the human eye or difficult to interpret clinically [47].

ML methods can integrate data from multiple sources (i.e., neuroimaging, genetics, clinical records, and wearable sensors) to provide a better overall view of the disease [48]. Such diverse data streams can be utilized by multimodal ML models to achieve improved diagnostic accuracy and provide more personalized insights [30]. A research by [32] supports this claim as it states, "Adopting multimodal learning within the medical domain facilitates the development of models aimed at enhancing the accuracy, predictability, and interpretability of medical diagnostics."

By employing ML algorithms to analyze high-dimensional datasets, it is possible to discover novel subtle patterns which can identify potential biomarkers for PD [49]. Building on this can lead to new diagnostics and therapeutic targets that can enhance clinician understanding of PD [50]. ML on clinical, genetic, and imaging data may further classify PD into subtypes [29]. This may enable more focused therapeutic approaches to distinct patient groups [51]. Research done by [32] shows that data-driven studies indicate that there is much that ML can do to help in the prediction of various subtypes of PD. It is also possible to analyze longitudinal data to track disease progression and predict future symptom development [29]. This information can inform treatment decisions and clinical trial design [52]. Personalized Treatment Plans and improved patient management are made possible through ML [53]. By considering individual patient characteristics and predicting treatment response, ML models can help personalize treatment plans [54]. This can lead to improved outcomes and reduced side effects [55]. Wearable sensor data combined with ML algorithms can enable remote monitoring of PD symptoms, allowing for early detection of changes in disease status and timely adjustments to treatment [34]. Hence, [40] in their research, recommend "the adoption of real-time and customized based devices with an advanced computing unit" for real-time PD diagnosis using image and sensory data. More importantly, insights gained from ML-driven biomarker discovery and disease modeling can inform the development of new drugs and therapeutic approaches for PD [29].

3. The contribution of XAI to the interpretability and trustworthiness of predictions in Parkinson's disease

XAI is crucial in enhancing the interpretability and trustworthiness of ML algorithms in the prediction of PD as it addresses the opaque nature of many ML models [28]. This is because traditional algorithms often make decisions or diagnoses without providing clear reasons and explaining the logic behind their reasoning, which then limits their acceptance by clinicians [32]. XAI techniques improve transparency by providing insight into the decision-making process, showing which features are most influential in predicting PD [27]. In essence, XAI helps clinicians understand the rationale behind its decisions. This is achieved through using methods such as SHAP, which quantify the contribution of each

feature to a prediction [31]. Additionally, techniques like LIME simplify complex models, giving a local understanding of predictions by showing how different features contribute and helping to validate the model's reasoning against expert knowledge.

XAI deepens our understanding of the mechanisms underlying PD, going beyond what traditional ML algorithms offer, for instance [27], XAI revealed the significance of features such as Difference Entropy and Joint Entropy in MRI scans, which show brain tissue degradation and structural changes crucial for early PD detection. This understanding of models focus on specific features enables healthcare personnel to better understand the relevant aspects of the disease, thus improving the differential diagnosis and assisting the creation of personalized treatment schemes.

In addition, explanatory XAI can be used to confirm the robustness of the ML-based prediction on PD by exposing the most significant features that the model relies on to make its decisions. This ensures that no spurious correlations or biases in the data generalized predictions [27]. This becomes especially important in health situations where life may hang in the balance. Through trust in ML models, the XAI would help advance the utilization of their clinical data by medical professionals, thereby improving trust that would eventually lead to the better reliability and usability of diagnostic tools for PD [31].

4. The main challenges and gaps of ML and XAI in Parkinson's disease prediction.

Current research into using ML and XAI for PD prediction faces multiple challenges and gaps. A significant problem that came up in reviewed studies was the issue of limited and imbalanced datasets. Various studies [30], [32], [36], [28] used data from a single modality such as MRI, speech, or gait instead of integrating types of data. A focus on single modalities limits the comprehensive view needed for accurate diagnosis as PD affects different aspects of a patient's health [40]. Furthermore, these datasets normally have more samples from healthy individuals than those with PD, sometimes leading to models biased toward the majority class [31]. The challenge with ML algorithms is that they are hard to explain, they have a "black box" nature which hinders clinical acceptance due to a lack of transparency and explainability [27]. Feature selection is also a challenge, as the feature selection methods in the models may not always align with the clinical understanding or be the most effective [32]. Furthermore, models may overfit the training data, thereby performing poorly on new, unseen data [56].

Another challenge is that XAI is that they have complex model explanations that are accessible to clinicians who may not be AI experts. There is a need to ensure that these explanations are not just technically sound but understandable to clinical staff [28]. Additionally, there is a lack of a standard, clinically trusted XAI framework that contributes to clinicians' hesitancy to adopt ML and XAI tools [31]. There is also a need for robust validation of XAI methods to make sure that they are sensitive and stable to small changes in input data [27]. Ensuring that ML models developed for PD perform consistently across different populations and clinical settings is essential. Factors such as variations in data acquisition protocols, patient demographics, and disease subtypes can affect model generalizability. The study by

[1], highlights that there is an issue of insufficient or inaccurate descriptions of methods or results in many studies, potentially hindering the reproducibility of results.

5. Limitations of model techniques

Data related limitations

The heterogeneity and availability of data present significant limitations. Available datasets often differ significantly, employing highly disparate modalities such as MRI, PET, SPECT, EEG, gait, handwriting, speech and clinical data, thus complicating integration and analysis [27]. Many good-quality, large datasets are often not publicly available [57]. Various studies [24], [25], [36], [34], [28], [31], [32], are also limited in sample size, making it difficult to generalize the results and increasing the likelihood of overfitting as limited data restricts effective training, especially with deep learning models [58]. Moreover, data imbalance is a problem where samples are heavily weighted towards one class (healthy controls) rather than the other (PD patients), and this can create biased models [59]. Furthermore, the performance of models can be negatively impacted by data quality issues, which can include noise, missing values, and inconsistencies in data collection [19]. Additionally, data collected in controlled settings may not accurately reflect real-world problems [1].

Methodological limitations

Feature selection i.e. identifying the most relevant features for accurate prediction, is difficult, requiring manual feature extraction, which is subjective and time-consuming [60]. Choosing the right model for specific tasks and data has proven to be difficult as some models are more suitable for certain types of data than others [1]. Training models on large datasets presents the risk of overfitting, where a model does not generalize to new data, although this can be mitigated somewhat by cross-validation and data augmentation [61]. Furthermore, several studies [30], [32], [36], [28], also exclusively focus on a single modality and overlook the fact that multiple modalities are necessary for the diagnosis of PD [62].

Clinical implementation limitations

This includes the absence of clinical validation of many models, meaning their potential to improve patient outcomes is unknown [49]. To bridge the gap between deep learning machine output and clinical use, XAI models can be incorporated in a clinical workflow to use traditional feature extraction, much like a neurologist would, however, deep learning models may not detect many of these same features seen by a neurologist in PD [4]. Healthcare professionals may also be reluctant to adopt new technologies due to psychological barriers, such as the endowment effect and status quo bias [63]. Another major problem is the absence of standardization, with many studies concentrating on single modalities rather than a multimodal methodology that is not feasible for clinical utility, whereby deep learning models usually identify PD using vectorized images rather than clinical inputs [64]. Clinicians also need interpretable models, that is, they need to know

why a model made a certain prediction to trust it [32]. Issues of regulatory approval and interoperability between disparate healthcare systems are also problems [4].

Another limitation is the subjectivity of clinical assessments. This can lead to varying levels of diagnostic accuracy and differential diagnosis. This makes it very difficult to differentiate PD from other Parkinsonian syndromes such as essential tremors [40]. Furthermore, variations in data collection protocols and in ethnicity may cause models developed on specific datasets or populations to not generalize to other groups [65]. These hurdles have to be overcome to realize the full benefits of ML and XAI in PD prediction and diagnosis.

6. The impact of ethical and privacy concerns the clinical application of ML and XAI in predicting PD and the strategies to mitigate these issues.

Ethics and privacy are primary considerations for the clinical application of ML and XAI in predicting PD [66]. One of the primary ethical issues that arise is "algorithmic bias". This is when ML algorithms reflect and amplify biases embedded in the data that the algorithms were trained on, resulting in false or misleading predictions for specific patient populations [67]. If a model is trained primarily on one demographic group, it may generalize to PD in other demographics [68]. This may lead to unequal access to care and less effective treatment for some patients [69]. Another ethical concern is that most ML models, particularly deep learning models, are not interpretable, functioning as black boxes [28]. Their decision-making processes are vague which makes it difficult to build trust among healthcare workers and patients [70]. The use of sensitive patient data is also of major concern as it brings up privacy and security issues particularly regarding compliance with privacy regulations [66], [71]. It is also important to consider the influence of medication on disease measures when defining Parkinson's disease states, as this can affect individual patients' clinical presentation [72].

To mitigate these privacy and ethical issues, we need to ensure data diversity and representativeness in training datasets to help reduce bias and improve the generalizability of ML models [30]. XAI techniques can be used to produce visual cues to better assess what features are involved in PD versus non-PD cases and to provide insight into ML model interpretability [70]. Data sharing should be encouraged to create larger cohorts of PD patients, but strict security guidelines must be followed to ensure data privacy [17][73][74]. Furthermore, addressing biases in data through careful data collection and pre-processing is essential to avoid unfair or inaccurate predictions [75]. A multidisciplinary environment combining clinical and computational expertise can foster the effective development and deployment of ML-based therapies such as adaptive Deep Brain Stimulation (aDBS) [69]. Regulatory bodies must develop guidelines for AI/ML-based medical devices, which include outlining good ML practices, setting guidelines for algorithm transparency, and establishing guidelines for real-world data collection and monitoring [76].

E. Implications of research findings

Theoretical Implications

The results show that ML can reshape PD diagnosis using some of the XAI methods such as SHAP and LIME for transparency and interpretability. ML can also reveal patterns in more complicated data, and it can also bring together multimodal data for personal medication and facilitate personalized medicine. This study adds to the growing body of knowledge about AI in health care, especially in exploring the comparative effectiveness of ML algorithms used in medical applications along with XAI in PD prediction. Leveraging XAI confronts a critical area in health care that leaves much to be desired in interpretability and trustworthiness of predictive models.

Practical Implications

This study highlights several practical ways in which ML and XAI are making a meaningful difference in predicting and managing PD. ML can capture the disease early enough so that timely intervention can enhance the quality of care and life for the patient, thus minimizing the burden of managing PD while freeing healthcare resources, especially where specialists are inadequate. Governments and policymakers must create a conducive environment for empowering ethical AI use, transparency, and compliance with healthcare standards. This must not be compromised in search for innovation, funding and collaboration in AI-driven healthcare research. Furthermore, XAI enhances early screening and risk assessment that is beneficial to health professionals, thereby improving patients' outcomes. Additionally, XAI-based prediction interpretation would further improve the trust toward these input's perceived performance, which encourages diagnostic process integration by clinicians. Integrating ML and XAI fosters in creating a need for collaboration with governments, health organizations, and AI scientists for responsible citizen integration and effective and equitable inclusion in the medical fraternity.

Limitations of the study

This SLR acknowledges several limitations, among which is the searching of only three databases with a restriction on publication dates (2020 to 2024). There is a possibility that there are more recent studies that were not included that were published after this search. These articles could have provided a deeper insight into ML and XAI in PD prediction. Only papers written in the English language were analyzed. This poses language bias and a risk that more substantive papers might have been written in other languages.

F. Conclusion and Future Works

This SLR explored the effectiveness of ML and XAI in predicting PD as compared to traditional methods. Combining ML with XAI in predicting PD can help identify the most critical features for accurate prediction and provide insights into the decision-making processes of the models, increasing trustworthiness and clinical applicability.

Despite these positive findings, the review highlights areas that need more attention, such as the act of incorporating more diverse and large datasets to address class imbalance and demographic representation and the integration of

Internet of Things (IoT) devices such as wearables and smart sensors to provide real-time, continuous patient data to improve model accuracy and early detection. Additionally, the adoption of trustworthy XAI techniques to establish ML models' transparency and interpretability and to gain trust from clinical practitioners and patients would greatly increase the chances of adoption in healthcare settings. Legal and ethical standards should be established, potentially leveraging blockchain technology, to ensure data integrity, patient privacy, and secure sharing of sensitive health information across stakeholders. Moreover, enabling collaborative work between data scientists, clinicians, and policymakers can ensure that AI solutions are well-suited to address clinical needs in practice. It is also essential to hold pilot studies and provide real-world validations as well as specific training for healthcare professionals to ensure easier incorporation and adoption of such superior technologies into clinical practice.

Furthermore, there should be a focus on integrating multiple data modalities for more accuracy while the development of robust real-time systems using edge computing will facilitate timely diagnosis and monitoring. Finally, there should be standardized reporting of methodologies and results along with comprehensive validation of results to promote the clinical translation of XAI and ML for PD diagnosis. Overall, this SLR concludes that ML and XAI show great promise for enhancing PD diagnosis.

G. References

[1] J. Mei, C. Desrosiers, and J. Frasnelli, "Machine Learning for the Diagnosis of Parkinson's Disease: A Review of Literature," May 06, 2021, *Frontiers Media S.A.* doi: 10.3389/fnagi.2021.633752.

[2] A. Rana, A. Dumka, R. Singh, M. K. Panda, N. Priyadarshi, and B. Twala, "Imperative Role of Machine Learning Algorithm for Detection of Parkinson's Disease: Review, Challenges and Recommendations," Aug. 01, 2022, *Multidisciplinary Digital Publishing Institute (MDPI)*. doi: 10.3390/diagnostics12082003.

[3] N. Salari, M. Kazeminia, H. Sagha, A. Daneshkhah, A. Ahmadi, and M. Mohammadi, "The performance of various machine learning methods for Parkinson's disease recognition: a systematic review," *Current Psychology*, vol. 42, no. 20, pp. 16637–16660, Jul. 2023, doi: 10.1007/s12144-022-02949-8.

[4] H. W. Loh *et al.*, "Application of deep learning models for automated identification of parkinson's disease: A review (2011–2021)," Nov. 01, 2021, *MDPI*. doi: 10.3390/s21217034.

[5] A. Rana, A. Dumka, R. Singh, M. K. Panda, and N. Priyadarshi, "A Computerized Analysis with Machine Learning Techniques for the Diagnosis of Parkinson's Disease: Past Studies and Future Perspectives," Nov. 01, 2022, *Multidisciplinary Digital Publishing Institute (MDPI)*. doi: 10.3390/diagnostics12112708.

[6] A. Balakrishnan, J. Medikonda, P. K. Namboothiri, and M. Natarajan, "Role of Wearable Sensors with Machine Learning Approaches in Gait Analysis for Parkinson's Disease Assessment: A Review," 2022, *Engineered Science Publisher*. doi: 10.30919/es8e622.

[7] Y. Miao, X. Lou, and H. Wu, "The Diagnosis of Parkinson's Disease Based on Gait, Speech Analysis and Machine Learning Techniques," in *Proceedings of the 2021 International Conference on Bioinformatics and Intelligent Computing, BIC 2021*, Association for Computing Machinery, Inc, Jan. 2021, pp. 358–371. doi: 10.1145/3448748.3448804.

[8] M. Shaban, "Deep Learning for Parkinson's Disease Diagnosis: A Short Survey," Mar. 01, 2023, *MDPI*. doi: 10.3390/computers12030058.

[9] N. W.C Mukura and B. Ndlovu, "Performance Evaluation of Artificial Intelligence in Decision Support System for Heart Disease Risk Prediction," no. Who 2018, pp. 83–93, 2023, doi: 10.46254/ap04.20230043.

[10] B. Ncube, M. Dzikiti, A. Nyoni, M. Ncube, and B. M. Ndlovu, "Effectiveness of Machine Learning algorithms in predicting Monkey Pox (Mpox): A Systematic Literature Review," 2024, doi: 10.46254/EU07.202400.

[11] K. Maguraushe, P. Ndayizigamiye, and T. Bokaba, "Trends and Developments in the Use of Machine Learning for Disaster Management: A Bibliometric Analysis," *IFIP Adv Inf Commun Technol*, vol. 698 AICT, pp. 92–104, 2024, doi: 10.1007/978-3-031-50192-0_9.

[12] L. Mpofu, B. Ndlovu, S. Dube, M. Muduva, F. Jacqueline, and K. Maguraushe, "Predictive Model for Hospital Readmission of Diabetic Patients," 2024, doi: 10.46254/af05.20240252.

[13] M. Sibindi *et al.*, "A Predictive Model for Personalized Healthcare Management for Patients with Chronic Diseases," 2024, doi: 10.46254/EU07.20240164.

[14] N. W. Chiramba, B. M. Ndlovu, and S. Dube, "Optimizing Antiretroviral Therapy (ART) Adherence through Predictive Analytics Using Machine Learning Techniques .," *IEOM*, vol. 2021, no. July, 2024, doi: 10.46254/SA05.20240195.

[15] V. Skaramagkas, A. Pentari, Z. Kefalopoulou, and M. Tsiknakis, "Multi-Modal Deep Learning Diagnosis of Parkinson's Disease - A Systematic Review," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 31, pp. 2399–2423, 2023, doi: 10.1109/TNSRE.2023.3277749.

[16] J. Martorell-Marugán, M. Chierici, S. Bandres-Ciga, G. Jurman, and P. Carmona-Sáez, "Machine Learning Applications in the Study of Parkinson's Disease: A Systematic Review," 2023, *Bentham Science Publishers*. doi: 10.2174/1574893618666230406085947.

[17] K. M. Giannakopoulou, I. Roussaki, and K. Demestichas, "Internet of Things Technologies and Machine Learning Methods for Parkinson's Disease Diagnosis, Monitoring and Management: A Systematic Review," Mar. 01, 2022, *MDPI*. doi: 10.3390/s22051799.

[18] A. M. Antoniadi *et al.*, "Current challenges and future opportunities for xai in machine learning-based clinical decision support systems: A systematic review," *Applied Sciences (Switzerland)*, vol. 11, no. 11, Jun. 2021, doi: 10.3390/app11115088.

[19] M. M. Alsaleh *et al.*, "Prediction of disease comorbidity using explainable artificial intelligence and machine learning techniques: A systematic review," Jul. 01, 2023, *Elsevier Ireland Ltd*. doi: 10.1016/j.ijmedinf.2023.105088.

[20] T. Kapuya, W. Kubiku, M. Dube, and T. Mukudu, "Ethical AI Frameworks : Balancing Privacy , Consent and Responsible Use," no. Landi, 2024, doi: 10.46254/EU07.20240168.

[21] S. S Band *et al.*, "Application of explainable artificial intelligence in medical health: A systematic review of interpretability methods," *Inform Med Unlocked*, vol. 40, Jan. 2023, doi: 10.1016/j.imu.2023.101286.
[22] K. Maguraushe and B. M. Ndlovu, "The use of smart technologies for enhancing palliative care: A systematic review.," *Digit Health*, vol. 10, p. 20552076241271836, 2024, doi: 10.1177/20552076241271835.
[23] D. Moher, A. Liberati, J. Tetzlaff, and D. G. Altman, "Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement," Aug. 08, 2009. doi: 10.1136/bmj.b2535.
[24] M. Abu Sayed *et al.*, "Parkinson's Disease Detection through Vocal Biomarkers and Advanced Machine Learning Algorithms," 2023, doi: 10.32996/jcsts.
[25] N. Fadavi and N. Fadavi, "Early Recognition of Parkinson's Disease Through Acoustic Analysis and Machine Learning," Jul. 2024, [Online]. Available: http://arxiv.org/abs/2407.16091
[26] Y. H. Park *et al.*, "Machine learning based risk prediction for Parkinson's disease with nationwide health screening data," *Sci Rep*, vol. 12, no. 1, Dec. 2022, doi: 10.1038/s41598-022-24105-9.
[27] S. Priyadharshini *et al.*, "A Comprehensive framework for Parkinson's disease diagnosis using explainable artificial intelligence empowered machine learning techniques," *Alexandria Engineering Journal*, vol. 107, pp. 568–582, Nov. 2024, doi: 10.1016/j.aej.2024.07.106.
[28] S. Saravanan, K. Ramkumar, K. Narasimhan, S. Vairavasundaram, K. Kotecha, and A. Abraham, "Explainable Artificial Intelligence (EXAI) Models for Early Prediction of Parkinson's Disease Based on Spiral and Wave Drawings," *IEEE Access*, vol. 11, pp. 68366–68378, 2023, doi: 10.1109/ACCESS.2023.3291406.
[29] S. Patil, S. Jaybhaye, S. Bokariya, P. Jain, S. Phapale, and T. Hande, "Parkinson's Disease Prediction System in Machine Learning," *ITM Web of Conferences*, vol. 56, p. 05002, 2023, doi: 10.1051/itmconf/20235605002.
[30] M. B. Makarious *et al.*, "Multi-modality machine learning predicting Parkinson's disease," *NPJ Parkinsons Dis*, vol. 8, no. 1, Dec. 2022, doi: 10.1038/s41531-022-00288-w.
[31] K. Shyamala and T. M. Navamani, "Design of an Efficient Prediction Model for Early Parkinson’s Disease Diagnosis," *IEEE Access*, 2024, doi: 10.1109/ACCESS.2024.3421302.
[32] V. Dentamaro, D. Impedovo, L. Musti, G. Pirlo, and P. Taurisano, "Enhancing early Parkinson's disease detection through multimodal deep learning and explainable AI: insights from the PPMI database," *Sci Rep*, vol. 14, no. 1, Dec. 2024, doi: 10.1038/s41598-024-70165-4.
[33] A. Kumar, S. Bind, A. K. Tiwari, and A. K. Sahani, "A Survey of Machine Learning Based Approaches for Parkinson Disease Prediction," 2022. [Online]. Available: https://www.researchgate.net/publication/362446752
[34] X. Yang, Q. Ye, G. Cai, Y. Wang, and G. Cai, "PD-ResNet for Classification of Parkinson's Disease From Gait," *IEEE J Transl Eng Health Med*, vol. 10, 2022, doi: 10.1109/JTEHM.2022.3180933.
[35] V. Ulagamuthalvi, "Identification of Parkinson's Disease Using Machine Learning Algorithms," *Biosci Biotechnol Res Commun*, vol. 13, no. 2, pp. 576–579, Jun. 2020, doi: 10.21786/bbrc/13.2/32.

[36] W. Wang, J. Lee, F. Harrou, and Y. Sun, "Early Detection of Parkinson's Disease Using Deep Learning and Machine Learning," *IEEE Access*, vol. 8, pp. 147635–147646, 2020, doi: 10.1109/ACCESS.2020.3016062.

[37] B. M. Ndlovu, B. Chipangura, and S. Singh, "Factors Influencing Quantified SelfTechnology Adoption in Monitoring Diabetes," in *Proceedings of Ninth International Congress on Information and Communication Technology*, X.-S. Yang, S. Sherratt, N. Dey, and A. Joshi, Eds., Singapore: Springer Nature Singapore, 2024, pp. 469–479.

[38] Yamada Y, Shinkawa K, Kobayashi M, Caggiano V, Nemoto M, Nemoto K, Arai T. Combining Multimodal Behavioral Data of Gait, Speech, and Drawing for Classification of Alzheimer's Disease and Mild Cognitive Impairment. J Alzheimers Dis. 2021;84(1):315-327. doi: 10.3233/JAD-210684.

[39] G. Airlangga, "Decoding Energy Usage Predictions: An Application of XAI Techniques for Enhanced Model Interpretability," Indones. J. Artif. Intell. Data Min., vol. 7, no. 2, p. 275, 2024, doi: 10.24014/ijaidm.v7i2.29041

[40] A. Rana, A. Dumka, R. Singh, M. Rashid, N. Ahmad, and M. K. Panda, "An Efficient Machine Learning Approach for Diagnosing Parkinson's Disease by Utilizing Voice Features," *Electronics (Switzerland)*, vol. 11, no. 22, Nov. 2022, doi: 10.3390/electronics11223782.

[41] M. Nazari *et al.*, "Explainable AI to improve acceptance of convolutional neural networks for automatic classification of dopamine transporter SPECT in the diagnosis of clinically uncertain parkinsonian syndromes," *Eur J Nucl Med Mol Imaging*, vol. 49, no. 4, pp. 1176–1186, Mar. 2022, doi: 10.1007/s00259-021-05569-9.

[42] A. Dadu *et al.*, "Identification and prediction of Parkinson's disease subtypes and progression using machine learning in two cohorts," *NPJ Parkinsons Dis*, vol. 8, no. 1, Dec. 2022, doi: 10.1038/s41531-022-00439-z.

[43] M. Nilashi *et al.*, "Predicting Parkinson's Disease Progression: Evaluation of Ensemble Methods in Machine Learning," *J Healthc Eng*, vol. 2022, 2022, doi: 10.1155/2022/2793361.

[44] J. M. Templeton, C. Poellabauer, and S. Schneider, "Classification of Parkinson's disease and its stages using machine learning," *Sci Rep*, vol. 12, no. 1, Dec. 2022, doi: 10.1038/s41598-022-18015-z.

[45] A. Govindu and S. Palwe, "Early detection of Parkinson's disease using machine learning," in *Procedia Computer Science*, Elsevier B.V., 2022, pp. 249–261. doi: 10.1016/j.procs.2023.01.007.

[46] L.-C. Chang *et al.*, "Machine learning approaches to identify Parkinson's disease using voice signal features."

[47] M. Pramanik, R. Pradhan, P. Nandy, A. K. Bhoi, and P. Barsocchi, "Machine learning methods with decision forests for parkinson's detection," *Applied Sciences (Switzerland)*, vol. 11, no. 2, pp. 1–25, Jan. 2021, doi: 10.3390/app11020581.

[48] N. Nahar, F. Ara, M. A. I. Neloy, A. Biswas, M. S. Hossain, and K. Andersson, "Feature Selection Based Machine Learning to Improve Prediction of Parkinson Disease," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in*

Bioinformatics), Springer Science and Business Media Deutschland GmbH, 2021, pp. 496–508. doi: 10.1007/978-3-030-86993-9_44.

[49] S. Aich *et al.*, "A supervised machine learning approach to detect the On/Off state in Parkinson's disease using wearable based gait signals," *Diagnostics*, vol. 10, no. 6, Jun. 2020, doi: 10.3390/diagnostics10060421.

[50] M. R. Salmanpour *et al.*, "Machine learning methods for optimal prediction of motor outcome in Parkinson's disease," *Physica Medica*, vol. 69, pp. 233–240, Jan. 2020, doi: 10.1016/j.ejmp.2019.12.022.

[51] A. Iyer *et al.*, "A machine learning method to process voice samples for identification of Parkinson's disease," *Sci Rep*, vol. 13, no. 1, Dec. 2023, doi: 10.1038/s41598-023-47568-w.

[52] P. Raundale, C. Thosar, and S. Rane, "Prediction of parkinson's disease and severity of the disease using machine learning and deep learning algorithm," in *2021 2nd International Conference for Emerging Technology, INCET 2021*, Institute of Electrical and Electronics Engineers Inc., May 2021. doi: 10.1109/INCET51464.2021.9456292.

[53] A. Moshkova, A. Samorodov, N. Voinova, E. Ivanova, and E. Fedotova, "Parkinson's Disease Detection by Using Machine Learning Algorithms and Hand Movement Signal from LeapMotion Sensor."

[54] C. Sotirakis *et al.*, "Identification of motor progression in Parkinson's disease using wearable sensors and machine learning," *NPJ Parkinsons Dis*, vol. 9, no. 1, Dec. 2023, doi: 10.1038/s41531-023-00581-2.

[55] F. Castelli Gattinara Di Zubiena *et al.*, "Machine Learning and Wearable Sensors for the Early Detection of Balance Disorders in Parkinson's Disease," *Sensors*, vol. 22, no. 24, Dec. 2022, doi: 10.3390/s22249903.

[56] D. Trabassi *et al.*, "Machine Learning Approach to Support the Detection of Parkinson's Disease in IMU-Based Gait Analysis," *Sensors*, vol. 22, no. 10, May 2022, doi: 10.3390/s22103700.

[57] A. Suppa *et al.*, "Voice in Parkinson's Disease: A Machine Learning Study," *Front Neurol*, vol. 13, Feb. 2022, doi: 10.3389/fneur.2022.831428.

[58] A. Li and C. Li, "Detecting Parkinson's Disease through Gait Measures Using Machine Learning," *Diagnostics*, vol. 12, no. 10, Oct. 2022, doi: 10.3390/diagnostics12102404.

[59] S. Moon *et al.*, "Classification of Parkinson's disease and essential tremor based on balance and gait characteristics from wearable motion sensors via machine learning techniques: A data-driven approach," *J Neuroeng Rehabil*, vol. 17, no. 1, Sep. 2020, doi: 10.1186/s12984-020-00756-5.

[60] F. Saeed *et al.*, "Enhancing Parkinson's Disease Prediction Using Machine Learning and Feature Selection Methods", doi: 10.32604/cmc.2022.xxxxxx.

[61] A. Boutet *et al.*, "Predicting optimal deep brain stimulation parameters for Parkinson's disease using functional MRI and machine learning," *Nat Commun*, vol. 12, no. 1, Dec. 2021, doi: 10.1038/s41467-021-23311-9.

[62] E. Pantaleo *et al.*, "A Machine Learning Approach to Parkinson's Disease Blood Transcriptomics," 2022, doi: 10.3390/genes.

[63] B. Mutunhu, B. Chipangura, and S. Singh, "An Exploration of Opportunities for Quantified-Self Technology in Diabetes Self-Care: A Systematic Literature

Review," *J Health Inform Afr*, vol. 11, no. 2, pp. 17–30, 2024, doi: 10.12856/JHIA-2024-v11-i2-481.

[64] A. Shcherbak, E. Kovalenko, and A. Somov, "Detection and Classification of Early Stages of Parkinson's Disease Through Wearable Sensors and Machine Learning," *IEEE Trans Instrum Meas*, vol. 72, 2023, doi: 10.1109/TIM.2023.3284944.

[65] H. Byeon, "Development of a depression in Parkinson's disease prediction model using machine learning," *World J Psychiatry*, vol. 10, no. 10, pp. 234–244, Oct. 2020, doi: 10.5498/wjp.v10.i10.234.

[66] B. Garcia Santa Cruz, A. Husch, and F. Hertel, "Machine learning models for diagnosis and prognosis of Parkinson's disease using brain imaging: general overview, main challenges, and future directions," 2023, *Frontiers Media SA*. doi: 10.3389/fnagi.2023.1216163.

[67] J. Zhang, "Mining imaging and clinical data with machine learning approaches for the diagnosis and early detection of Parkinson's disease," Dec. 01, 2022, *Nature Research*. doi: 10.1038/s41531-021-00266-8.

[68] M. Shen, P. Mortezaagha, and A. Rahgozar, "Explainable Artificial Intelligence to Diagnose Early Parkinson's Disease via Voice Analysis," Sep. 30, 2024. doi: 10.1101/2024.09.29.24314580.

[69] A. M. Oliveira, L. Coelho, E. Carvalho, M. J. Ferreira-Pinto, R. Vaz, and P. Aguiar, "Machine learning for adaptive deep brain stimulation in Parkinson's disease: closing the loop," Nov. 01, 2023, *Springer Science and Business Media Deutschland GmbH*. doi: 10.1007/s00415-023-11873-1.

[70] P. R. Magesh, R. D. Myloth, and R. J. Tom, "An Explainable Machine Learning Model for Early Detection of Parkinson's Disease using LIME on DaTscan Imagery," Aug. 2020, [Online]. Available: http://arxiv.org/abs/2008.00238

[71] K. Maguraushe, A. da Veiga, and N. Martins, "A personal information privacy perceptions model for university students," *Information Security Journal: A Global Perspective*, vol. 33, no. 4, pp. 394–424, Jul. 2024, doi: 10.1080/19393555.2024.2329554.

[72] K. A. Severson *et al.*, "Discovery of Parkinson's disease states and disease progression modelling: a longitudinal data study using machine learning," *Lancet Digit Health*, vol. 3, no. 9, pp. e555–e564, Sep. 2021, doi: 10.1016/S2589-7500(21)00101-1.

[73] Mutunhu, B., Chipangura, B., & Singh, S. (2024b). Towards a quantified-self technology conceptual framework for monitoring diabetes. *South African Journal for Science and Technology*, *43*(1), 69–84. https://doi.org/10.36303/SATNT.2024.43.1.970

[74] K. Maguraushe, "Development of a diagnostic instrument and privacy model for student personal information privacy perceptions at a Zimbabwean university Kudakwashe Maguraushe," 2021.

[75] A. M. Maitín, A. J. García-Tejedor, and J. P. R. Muñoz, "Machine learning approaches for detecting parkinson's disease from eeg analysis: A systematic review," Dec. 01, 2020, *MDPI AG*. doi: 10.3390/app10238662.

[76] A. S. Chandrabhatla, I. J. Pomeraniec, and A. Ksendzovsky, "Co-evolution of machine learning and digital technologies to improve monitoring of

Parkinson's disease motor symptoms," Dec. 01, 2022, *Nature Research*. doi: 10.1038/s41746-022-00568-y.