Scientific Reports

https://doi.org/10.1038/s41598-026-62606-z

Article in Press

Quantum-classical hybrid learning framework for Parkinson's disease prediction using acoustic and clinical features

Received: 15 December 2025

Accepted: 13 July 2026

Published online: 23 July 2026

Cite this article as: Li J., Rao N.R., Kumar C.H.N.S. *et al.* Quantum-classical hybrid learning framework for Parkinson's disease prediction using acoustic and clinical features. *Sci Rep* (2026). https://doi.org/10.1038/s41598-026-62606-z

Jiahua Li, N. Rama Rao, C. H. N. Santhosh Kumar, G. Madhukar Rao, Pranto Halder, Shakti Singh & Doğan Keskin

We are providing an unedited version of this manuscript to give early access to its findings. Before final publication, the manuscript will undergo further editing. Please note there may be errors present which affect the content, and all legal disclaimers apply.

If this paper is publishing under a Transparent Peer Review model then Peer Review reports will publish with the final article.

ARTICLE IN PRESS

© The Author(s) 2026. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

ARTICLE IN PRESS

Quantum-Classical Hybrid Learning Framework for Parkinson's Disease Prediction Using Acoustic and Clinical Features

Jiahua Li[1], N. Rama Rao[2], CH N Santhosh Kumar[3], G. Madhukar Rao[4], Pranto Halder[5], Shakti Singh[6], Doğan Keskin[7]*

[1]School of Computer Science and Technology, Zhuhai College of Science, No. 8 Anji East Road, Sanzhao Town, Jinwan District, Zhuhai, Guangdong, 519041, China.

bsljh0813@163.com

[2]Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Aziz Nagar, Hyderabad, Telangana, 500075, India. **ramarao.n@klh.edu.in**

[3]Department of Computer Science and Engineering, Anurag Engineering College, Kodad, Telangana, India., **santhosh.ph10@gmail.com**

[4]Department of Computer Science and Engineering, Koneru Lakshmaiah Educational Foundation, Aziz Nagar, Hyderabad, Telangana, 500075, India. **madhusw511@klh.edu.in**

[5]Dept. of Computer Science and Engineering, Faculty of Science and Engineering, Teesta University, Rangpur-5404, Bangladesh.

prantohalder117@gmail.com

[6]Department of Computer Science and Engineering, Lovley Professional University Phagwara, Punjab; **shaktisingh789880@gmail.com**

[7]Department of Computer Science, Turkish German University, Istanbul, Turkey;

Faculty of Computing & Information Technology, University of Yemen, Sana'a, Yemen.

dogankz6777@gmail.com *(Corresponding Author)*

ARTICLE IN PRESS

Abstract

Parkinson's disease (PD) is a progressive neurodegenerative disorder affecting millions worldwide, where early and accurate diagnosis plays a critical role in improving patient outcomes. While classical machine learning methods have shown strong performance in PD classification using clinical and acoustic data, they remain limited in capturing complex nonlinear feature interactions in high-dimensional spaces. This study explores the application of quantum machine learning (QML) techniques for PD prediction, focusing on Variational Quantum Classifiers (VQC) and Quantum Support Vector Machines (QSVM), and compares their performance with strong classical baselines including Logistic Regression, SVM, Random Forest, Grid Search–optimized models, and XGBoost. Experiments were conducted using stratified train–test splits and validated through cross-validation and statistical testing. Results show that classical models achieve competitive performance, with XGBoost reaching 91.9%

ARTICLE IN PRESS

accuracy. However, quantum models demonstrate slightly improved performance, with the VQC using EfficientSU2 achieving the highest accuracy of 93.1%, followed by the Real Amplitudes VQC (92.6%) and QSVM (92.2%). These improvements are consistent across evaluation metrics including F1-score and recall. Statistical analysis using 10-fold cross-validation indicates that the VQC (EfficientSU2) shows a significant improvement over classical baselines ($p = 0.031$), while other quantum models do not achieve statistical significance over XGBoost ($p > 0.05$). McNemar's test on the independent test set confirms no significant difference in final prediction distributions between quantum and classical models. Overall, the findings suggest that quantum machine learning models are competitive with state-of-the-art classical approaches, offering marginal but consistent performance gains. This study highlights the potential of quantum-enhanced learning frameworks for biomedical classification tasks, while also emphasizing that their advantages remain incremental rather than definitive under current simulation-based implementations.

Keywords: Parameterized Quantum Models, Quantum-Enhanced Learning Methods, Quantum Information Processing, Neurodegenerative Movement Disorder

1. Introduction

Millions of people worldwide are affected by Parkinson's disease (PD), a progressive neurodegenerative disorder characterized primarily by the degeneration of dopaminergic neurons in the substantia nigra, leading to dopamine deficiency that significantly impairs motor control and coordination [1]. The most commonly observed motor symptoms include tremors, bradykinesia (slowness of movement), muscle rigidity, and postural instability. In addition to these motor impairments, patients frequently experience a wide range of non-motor symptoms such as cognitive decline, sleep disturbances, mood disorders, depression, and anxiety, which collectively contribute to a substantial reduction in quality of life.

Although the exact etiology of Parkinson's disease remains unclear, current evidence suggests that its onset and progression are influenced by a complex interplay of genetic predisposition, environmental exposures, and biological factors [2]. Due to the absence of a definitive diagnostic test, PD is primarily diagnosed through clinical evaluation, neurological examination, and longitudinal symptom assessment. Existing therapeutic approaches, including pharmacological treatments such as levodopa, surgical interventions like deep brain stimulation, and supportive therapies such as physiotherapy, are primarily symptomatic in nature and do not provide a cure. Nevertheless, evidence indicates that early diagnosis and timely intervention can significantly improve disease management and patient outcomes [3].

The global prevalence of Parkinson's disease is increasing at an alarming rate, driven by aging populations, improved diagnostic capabilities, and greater awareness and reporting of neurodegenerative disorders. Recent epidemiological projections suggest that the number of individuals affected by PD may double over the next two decades [4]. This rapid rise places a considerable burden on healthcare systems, caregivers, and socioeconomic structures worldwide, highlighting the urgent need for more effective diagnostic, monitoring, and therapeutic strategies.

In clinical practice, much of the focus has traditionally been placed on motor symptoms, while non-motor manifestations have often been underdiagnosed or overlooked. However, growing evidence emphasizes that non-motor symptoms such as cognitive impairment, apathy, sleep dysfunction, and psychiatric disturbances are equally critical in determining disease progression

ARTICLE IN PRESS

and patient well-being [5]. These symptoms frequently complicate diagnosis and treatment planning, underscoring the need for comprehensive and holistic disease management approaches that integrate both motor and non-motor dimensions of PD. Recent advancements in digital health technologies, artificial intelligence, wearable sensors, and biomedical imaging are transforming the landscape of Parkinson's disease diagnosis and management. These innovations enable continuous monitoring of patient behavior, real-time symptom tracking, and the development of personalized treatment strategies [6]. As a result, healthcare is gradually shifting from traditional episodic assessments toward more dynamic, data-driven, and patient-centric care models [7]. In addition, emerging research highlights the potential role of the gut–brain axis and the human microbiome in the pathophysiology of Parkinson's disease [8]. Increasing evidence suggests that gastrointestinal dysfunction and microbiota alterations may precede classical motor symptoms by several years, indicating their potential as early biomarkers for disease onset [9]. Understanding the bidirectional interaction between the central nervous system and systemic biological processes may open new avenues for early detection, preventive strategies, and minimally invasive therapeutic interventions. Overall, Parkinson's disease represents a complex, multifactorial neurological condition that requires interdisciplinary research efforts integrating neuroscience, biomedical engineering, and artificial intelligence. Advancements in these domains hold significant promise for improving early diagnosis, enhancing disease monitoring, and ultimately improving the quality of life for affected individuals.

1.1. Quantum Machine Learning

Quantum Machine Learning (QML), an emerging interdisciplinary field at the intersection of quantum computing and machine learning, holds significant promise for transforming computational efficiency and complex problem-solving capabilities [10]. Healthcare is one of the most data-intensive domains, generating massive volumes of heterogeneous information through clinical diagnostics, imaging systems, electronic health records, and biomedical monitoring devices. Efficient processing and interpretation of such large-scale medical data are essential for understanding biological processes and improving patient-centered care.

Quantum computing introduces a fundamentally new computational paradigm by leveraging quantum bits (qubits), which operate based on the principles of superposition and entanglement. These properties enable quantum systems to perform parallel computations, potentially achieving exponential speedups over classical computing methods [11]. As a result, quantum computing offers a promising solution to several limitations faced by traditional machine learning algorithms, particularly in handling high-dimensional and complex datasets.

Unlike classical systems, quantum computers utilize qubits that can exist in multiple states simultaneously, enabling significantly faster and more efficient processing of information [12]. This capability has led to increasing interest in integrating quantum computing with machine learning to enhance predictive modeling, optimization, and classification tasks [13]. In this context, various quantum-enhanced machine learning techniques have been proposed, including Quantum Support Vector Classifier (QSVC), Variational Quantum Circuits (VQC), Support

ARTICLE IN PRESS

Vector Classifier with quantum kernels, and hybrid models incorporating quantum-augmented feature spaces [14]. These quantum-based approaches are typically evaluated against traditional Support Vector Machine (SVM) models to assess their effectiveness in handling complex experimental datasets. Early findings suggest that quantum-enhanced models may offer improved classification performance and better scalability for certain high-dimensional problems, although practical implementation remains an active area of research. In the healthcare domain, such advancements hold particular relevance for improving diagnostic accuracy and personalizing treatment strategies. One potential application involves optimizing psychotropic medication regimens for individuals suffering from severe depressive disorders (SDD). By integrating quantum machine learning techniques with clinical and genomic data, it becomes possible to enhance predictive analysis and support more precise treatment recommendations. Furthermore, this research explores the relationship between severe depressive disorder, conventional psychotropic treatments, and cytochrome-based pharmacogenetic profiling, aiming to better understand individual variations in drug response. The integration of quantum machine learning into this framework provides an opportunity to model complex biological and psychological interactions more effectively than classical approaches.

1.2. Roles of QML in Disease Detection

Quantum Machine Learning (QML) is emerging as a transformative paradigm in disease detection, addressing the limitations of classical computing in processing large-scale and highly complex medical datasets. Modern medical imaging modalities such as Magnetic Resonance Imaging (MRI), Computed Tomography (CT), and Positron Emission Tomography (PET) generate massive volumes of high-dimensional data, requiring advanced computational techniques for efficient analysis and interpretation [17]. QML, through quantum-enhanced supervised and hybrid algorithms, enables faster and more accurate identification of patterns and anomalies within such imaging data, thereby supporting early diagnosis of both neurological and systemic diseases. Early detection plays a critical role in reducing mortality rates, lowering healthcare costs, and significantly improving patient outcomes [18]. Beyond imaging, QML has shown strong potential in genomic and proteomic analysis, where it can process vast biological datasets to identify disease-associated genetic markers. By efficiently analyzing mutations, gene expressions, and molecular interactions, QML enhances genome-wide association studies (GWAS) and accelerates the discovery of clinically relevant biomarkers [19]. This capability supports the advancement of precision medicine by enabling treatment strategies that are tailored to an individual's genetic profile. As a result, quantum-enhanced approaches contribute to a deeper understanding of disease mechanisms and facilitate the development of highly personalized therapeutic interventions [20].

Another important advantage of QML lies in its ability to integrate and analyze heterogeneous healthcare data sources. By combining information from wearable health monitoring devices, electronic health records, and real-time physiological data streams, QML-based systems can construct highly accurate predictive models for early detection of chronic conditions such as

ARTICLE IN PRESS

diabetes, cardiovascular diseases, and hypertension, as well as emerging infectious diseases [21]. This multi-source data fusion enables a more comprehensive and proactive healthcare framework, shifting the focus from reactive treatment to preventive diagnosis. The synergy between quantum computing and machine learning enhances both computational speed and predictive accuracy, enabling the development of intelligent healthcare systems capable of continuous monitoring and early intervention. Although QML is still in its developmental phase, its potential to revolutionize disease detection, monitoring, and treatment strategies is increasingly evident, particularly in enabling proactive and personalized healthcare ecosystems [22]. In this context, the present research introduces a novel framework that integrates multimodal clinical, genetic, and behavioral data with Quantum Machine Learning (QML) techniques to improve the early detection, classification, and personalized treatment of Parkinson's disease. Unlike conventional machine learning approaches, this study demonstrates the effectiveness of hybrid quantum–classical models, particularly Quantum Support Vector Classifiers (QSVCs), in capturing complex, high-dimensional patterns within medical datasets. Furthermore, this work distinguishes itself by incorporating explainable AI (XAI) techniques to enhance interpretability and clinical trust. The proposed approach identifies previously underexplored biomarkers, such as vocal amplitude variations and cortical thinning patterns, which are critical in Parkinson's disease progression. By bridging advanced quantum analytics with clinically meaningful interpretability, this research narrows the gap between computational intelligence and real-world medical decision-making, paving the way for next-generation intelligent healthcare systems.

1.3 Motivation and Contribution

Parkinson's disease (PD) significantly deteriorates the quality of life of affected individuals due to its progressive neurodegenerative nature, primarily impacting motor control and cognitive functions. Early detection and accurate classification are essential for timely intervention, improved disease management, and enhanced patient outcomes. Traditionally, machine learning techniques have demonstrated effectiveness in processing medical datasets for Parkinson's disease detection. However, these classical approaches often face limitations in scalability, computational efficiency, and performance when applied to large-scale, high-dimensional clinical datasets. To overcome these challenges, Quantum Machine Learning (QML), an emerging paradigm inspired by quantum physics principles, has been introduced as a promising alternative. QML leverages quantum computational capabilities to enhance processing speed, optimize feature representation, and improve classification accuracy in complex datasets. In this study, clinical data-driven QML techniques are explored to evaluate their effectiveness in predicting and classifying Parkinson's disease. By integrating quantum-enhanced learning models, the proposed approach aims to improve both diagnostic precision and computational efficiency, thereby offering a potential advancement in the medical diagnosis of complex neurological disorders.

The research framework was developed through collaborative efforts involving the design of methodology, data preprocessing, model development, and integration of quantum machine learning algorithms for classification tasks. The study emphasizes the application of quantum-

ARTICLE IN PRESS

enhanced classifiers to analyze clinical Parkinson's datasets and assess their performance against traditional machine learning approaches.

The key contributions of this work include:

- Development and implementation of quantum machine learning classifiers with a focus on algorithmic optimization and quantum-inspired computational modeling.
- Execution of feature selection, data preprocessing, and systematic experimental evaluation of proposed models.
- Comprehensive analysis of results to assess both clinical relevance and statistical significance of classification performance.

Overall, the findings suggest that quantum machine learning holds strong potential for improving diagnostic frameworks in neurology, particularly for complex disorders such as Parkinson's disease. The proposed approach contributes toward the development of more advanced, accurate, and computationally efficient diagnostic systems in modern healthcare.

The remainder of this paper is organized as follows. Section 2 presents the dataset description and ethical considerations, along with preprocessing and feature engineering steps. Section 3 details the proposed methodology, including classical baseline models, quantum-enhanced learning frameworks, model architectures, and experimental setup. Section 4 describes the evaluation strategy, including performance metrics, cross-validation procedures, and statistical significance tests used for model comparison. Section 5 presents and discusses the experimental results, highlighting performance comparisons between classical and quantum approaches along with interpretability analysis. Section 6 outlines the limitations of the study and suggests directions for future research. Finally, Section 7 concludes the paper by summarizing key findings and contributions.

ARTICLE IN PRESS

2. Related work

Parkinson's disease (PD) is a progressive neurodegenerative disorder that affects millions of individuals globally and continues to pose significant challenges to healthcare systems due to its complex and multifactorial nature. The disease primarily impacts motor control, leading to symptoms such as tremors, bradykinesia, rigidity, and postural instability, while also causing a range of non-motor impairments including cognitive decline, mood disorders, and sleep disturbances. Given its irreversible progression, early diagnosis plays a crucial role in improving patient outcomes, slowing disease advancement, and enabling more effective therapeutic interventions. However, traditional diagnostic approaches based on clinical examination and neuroimaging often suffer from subjectivity, variability, and high resource dependency.

To address these limitations, computational intelligence techniques, particularly machine learning (ML) and deep learning (DL), have been increasingly adopted to analyze clinical, behavioral, and biomedical data for Parkinson's disease detection and progression analysis. These methods offer enhanced capability in handling complex datasets and identifying subtle patterns that may not be easily observable through conventional diagnostic techniques. In recent

ARTICLE IN PRESS

research, convolutional neural networks (CNNs) have been widely applied to neuroimaging data, particularly dopamine transporter scans, to improve diagnostic accuracy. A notable study [23] demonstrated the effectiveness of CNN-based architectures in classifying Parkinson's disease with an accuracy of approximately 92%. Importantly, the model also reduced false negative rates, which is critical for early-stage diagnosis where missed detection can significantly affect treatment outcomes. This work highlighted the limitations of manual imaging interpretation and reinforced the role of artificial intelligence in augmenting clinical decision-making processes.Beyond classification tasks, explainable artificial intelligence (XAI) has also emerged as a significant advancement in Parkinson's disease research. Study [24] developed an interpretable AI framework aimed at predicting disease progression while simultaneously identifying clinically relevant biomarkers. The model successfully uncovered hidden indicators such as cortical thinning and white matter abnormalities, which are often overlooked in traditional assessments. By improving interpretability, this approach bridges the gap between predictive analytics and clinical usability, enabling physicians to make more informed decisions.

In addition to imaging-based approaches, speech analysis has gained considerable attention as a non-invasive diagnostic tool. Research [25] investigated subtle changes in vocal characteristics, including tremors and reduced vocal amplitude, for early detection of Parkinson's disease. By employing hybrid quantum-classical Support Vector Machines (SVMs), the study achieved a 15% improvement in sensitivity compared to conventional methods. This demonstrates the potential of speech-based biomarkers as accessible and cost-effective screening tools for early diagnosis. Similarly, gait analysis has been explored as an important indicator of motor dysfunction in Parkinson's disease. Study [26] focused on analyzing gait parameters such as stride length, variability, and postural instability. Machine learning models were able to identify nearly 90% of disease progression patterns, demonstrating the effectiveness of movement-based data in clinical evaluation. Furthermore, the study emphasized the role of wearable sensors in enabling continuous, real-time monitoring of patients, offering a scalable and affordable solution for long-term disease management.

Genetic studies have also contributed significantly to understanding Parkinson's disease etiology. Research [27] focused on genetic mutations such as SNCA and LRRK2 and identified more than 30 genetic loci associated with disease susceptibility. These findings provide valuable insights into personalized medicine approaches, enabling the development of targeted therapies based on individual genetic profiles. This reinforces the importance of genetic predisposition in disease onset and progression.

Multi-modal data integration has further enhanced diagnostic accuracy in Parkinson's disease research. Study [28] combined neuroimaging data, blood biomarkers, and cerebrospinal fluid analysis to develop a predictive model achieving an Area Under the Curve (AUC) of 0.95. This multi-source integration approach demonstrated the importance of combining heterogeneous datasets to obtain a more comprehensive understanding of disease pathology and improve early

ARTICLE IN PRESS

detection capabilities.

With the emergence of advanced computational paradigms, quantum-enhanced machine learning has introduced new possibilities for handling high-dimensional biomedical data. Study [29] demonstrated that quantum-enhanced Support Vector Machines (QSVMs) are capable of efficiently processing complex datasets involving motor symptoms and neuroimaging data. The quantum kernel approach improved classification accuracy by capturing intricate feature dependencies that are difficult for classical models to represent effectively. Further advancements in quantum computing applications were reported in [30], where quantum-enhanced SVMs achieved a 20% improvement in processing speed compared to traditional SVM models. This enhancement addresses critical scalability challenges associated with large clinical datasets, making quantum machine learning a promising candidate for future healthcare applications.In terms of therapeutic innovations, non-invasive treatment approaches such as focused ultrasound have been explored for symptom management in Parkinson's disease. Study [31] demonstrated that focused ultrasound can effectively reduce tremors and rigidity, with approximately 75% of patients reporting significant symptom improvement and minimal side effects. This method presents a promising alternative to invasive procedures such as deep brain stimulation.

In addition, growing evidence suggests a strong connection between gut microbiota and Parkinson's disease progression. Research [32] highlighted the role of microbiome alterations in influencing disease severity and progression. Probiotic-based interventions targeting specific microbial strains have shown potential in slowing motor symptom development, indicating a shift toward microbiome-based therapeutic strategies.

During the COVID-19 pandemic, telemedicine emerged as a crucial tool for maintaining continuity of care in Parkinson's disease management. Study [33] reported that approximately 80% of patients found virtual consultations to be as effective as in-person visits. This highlights the importance of digital healthcare platforms in improving accessibility, especially for patients with mobility limitations. Similarly, mobile health technologies have contributed significantly to patient monitoring and personalized treatment. Study [34] developed a smartphone-based application capable of tracking tremor intensity and medication response in real time. The system enabled data-driven adjustments to treatment plans, demonstrating the potential of digital tools in enhancing patient-centered care. From a therapeutic perspective, pharmacological advancements targeting genetic and molecular mechanisms have shown promising outcomes. Study [35] identified LRRK2 inhibitors as potential neuroprotective agents, particularly for genetically predisposed patients, with 60% of participants showing slowed disease progression in clinical trials. Furthermore, research [36] on alpha-synuclein-targeting monoclonal antibodies has demonstrated early success in reducing protein aggregation, with ongoing Phase III clinical trials aiming to validate their long-term effectiveness.

Table 1 summarizes these studies by categorizing methodologies, datasets, performance metrics,

ARTICLE IN PRESS

and evaluation techniques used in Parkinson's disease research. The table provides a structured comparison of traditional machine learning models, deep learning frameworks, quantum-enhanced approaches, and emerging therapeutic strategies. Performance indicators such as accuracy, recall, F1-score, and AUC are used to evaluate the effectiveness of each method. This structured overview highlights the interdisciplinary nature of Parkinson's disease research, spanning domains such as neuroimaging, wearable technology, genomics, speech analysis, and quantum computing.

Overall, the reviewed literature demonstrates a clear evolution from traditional diagnostic methods to advanced AI-driven and quantum-enhanced frameworks. Despite significant progress, challenges remain in terms of data heterogeneity, model interpretability, and clinical deployment. Future research is expected to focus on developing unified multi-modal frameworks, integrating explainable AI techniques, and leveraging quantum machine learning to further enhance diagnostic accuracy and personalized treatment strategies for Parkinson's disease.

Table 1: Disease Detection using Quantum Machine Learning Classifiers

Ref.	**Year**	**Method/Tools used**	**Dataset**	**Implementation**	**Performance Evaluation**		
					Accuracy	**F1-Score**	**Recall**
[23]	2022	CNN on dopamine transporter scans	Neuroimaging DAT scans	✓	(92%)	✓	X
[24]	2023	Explainable AI. neuroimaging biomarkers	MRI cortical & white matter data	✓	X	✓	X
[25]	2021	Hybrid quantum-classical SVM on speech	Speech signal recordings	✓	✓ (+15%)	✓	X
[26]	2023	ML on gait analysis, wearable tech	Wearable sensor gait data	✓	(90%)	✓	X
[27]	2020	Meta-analysis, GWAS for PD loci	Genetic studies (SNCA, LRRK2,	X	X	X	X

ARTICLE IN PRESS

			GWAS)				
[28]	2024	Multi-modal predictive model (AUC 0.95)	Imaging, blood, and CSF biomarker data	✓	✓ (AUC 0.95)	✓	✓
[29]	2019	Quantum-enhanced SVM	Motor symptoms + neuroimaging data	✓	✓	✓	X
[30]	2023	Quantum-enhanced SVMs for scalability	Clinical datasets	✓	(20% faster)	✓	X
[31]	2022	Focused ultrasound therapy	Clinical trial participants	✓	X	✓ (75%)	X
[32]	2024	Gut microbiota, probiotic treatment	Microbiome and clinical symptom data	✓	X	✓	X
[33]	2020	Telemedicine during COVID-19	Patient telehealth surveys	✓	(80%)	✓	X
[34]	2023	Smartphone app for real-time tracking	App-based patient usage data	✓	X	✓	X
[35]	2022	LRRK2 inhibitors (neuroprotective drugs)	Genetic PD cohort (clinical trials)	✓	(60%)	✓	X
[36]	2024	Alpha-synuclein therapies (phase 3)	Early-stage PD trial population	X	X	X	X

ARTICLE IN PRESS

3. Proposed Methodology

This study proposes an innovative framework based on Quantum Machine Learning (QML) classifiers for the prediction and classification of Parkinson's disease (PD). Although traditional machine learning techniques have demonstrated considerable effectiveness in PD detection, they often encounter limitations related to high-dimensional feature spaces, computational complexity, and suboptimal scalability when applied to large biomedical datasets. To address

ARTICLE IN PRESS

these challenges, the integration of quantum computing is explored with the aim of improving both classification accuracy and computational efficiency, particularly for complex and high-dimensional medical data. The proposed methodology consists of several key stages, including data preprocessing, feature extraction, quantum state encoding, quantum classifier training, and performance evaluation. Initially, publicly available Parkinson's disease datasets are preprocessed to extract clinically relevant features such as speech patterns, gait characteristics, and biological signal indicators. These features capture both motor and non-motor aspects of the disease and serve as the input for subsequent quantum processing. In the next stage, classical feature vectors are transformed into quantum representations using encoding techniques such as basis encoding and amplitude encoding. This step enables the mapping of biomedical data into quantum states, allowing quantum algorithms to process information in a fundamentally different computational space. Following encoding, quantum classifiers are trained to learn complex patterns within the data and perform disease classification tasks. The final stage involves evaluating the performance of the proposed QML models using standard classification metrics. The objective is to assess the effectiveness of quantum-enhanced learning in improving diagnostic precision and computational efficiency compared to classical approaches. This study highlights the potential of Quantum Machine Learning in transforming Parkinson's disease diagnosis by improving model scalability, adaptability, and predictive performaTnce. The findings suggest that integrating quantum computing with biomedical data analysis can contribute to the development of more advanced and efficient neurological diagnostic systems, ultimately supporting more accurate and timely disease detection.

3.1 Dataset & Description

The dataset used in this study is the **Parkinson's Telemonitoring Dataset**, which is publicly available through the UCI Machine Learning Repository. The dataset was developed by **Athanasios Tsanas** and **Max A. Little** from the University of Oxford in collaboration with ten medical centers across the United States and Intel Corporation. It was collected as part of a six-month telemonitoring study aimed at remotely monitoring the progression of Parkinson's disease using speech signals recorded in patients' home environments. The dataset comprises **5,875 voice recordings** obtained from **42 individuals diagnosed with early-stage Parkinson's disease**. Each participant contributed approximately 200 sustained phonation recordings captured using a telemonitoring device. Every record corresponds to a single voice recording and contains demographic information, clinical assessments, and biomedical acoustic measurements. The demographic variables include **subject identifier, age, gender**, and **time elapsed since recruitment**, while the clinical outcomes include the **Motor Unified Parkinson's Disease Rating Scale (Motor UPDRS)** and **Total Unified Parkinson's Disease Rating Scale (Total UPDRS)** scores, which quantify disease severity. In addition to the clinical variables, the dataset contains **16 biomedical voice features** extracted from each speech recording. These features characterize frequency perturbation (Jitter-related measures), amplitude perturbation (Shimmer-related measures), harmonicity (Noise-to-Harmonic Ratio and Harmonic-to-Noise Ratio), and nonlinear speech dynamics, including **Recurrence Period Density Entropy (RPDE)**, **Detrended Fluctuation Analysis (DFA)**, and **Pitch Period Entropy (PPE)**. These acoustic biomarkers have been widely reported as reliable indicators of speech impairment associated with Parkinson's disease.

ARTICLE IN PRESS

The primary objective of the dataset is to predict **Motor UPDRS** and **Total UPDRS** scores from the extracted acoustic voice features, thereby enabling objective assessment of disease severity through non-invasive speech analysis. Owing to its comprehensive set of clinically relevant voice biomarkers and longitudinal recordings, the dataset provides an appropriate benchmark for evaluating both classical machine learning and quantum machine learning models for Parkinson's disease assessment [37].

Table 2. Comprehensive Description of Clinical, Demographic, and Biomedical Features in Parkinson's Disease Dataset

Category	Attribute(s)	Description	Data Type
Dataset Source	Parkinson's Telemonitoring Dataset	Publicly available dataset developed by Athanasios Tsanas and Max A. Little (University of Oxford) in collaboration with 10 U.S. medical centers and Intel Corporation	-
Participants	Subjects	42 individuals diagnosed with early-stage Parkinson's disease	Numerical
Study Duration	Monitoring Period	Six-month telemonitoring trial with speech recordings collected in patients' homes	-
Total Samples	Voice Recordings	5,875 sustained phonation recordings (approximately 200 recordings per subject)	Numerical
Recording Format	Data Format	ASCII CSV, where each row represents a single voice recording	-
Patient Information	Subject	Unique identifier assigned to each participant	Integer
	Age	Age of the participant (years)	Numerical
	Sex	Gender of the participant (0 = Male, 1 = Female)	Binary
	Test Time	Time elapsed since recruitment into the study (days)	Numerical
Clinical Measures	Motor_UPDRS	Clinician-assessed motor Unified Parkinson's Disease Rating Scale score	Numerical
	Total_UPDRS	Clinician-assessed total Unified Parkinson's Disease Rating Scale score	Numerical
Frequency Perturbation Features	Jitter (%), Jitter (Abs), Jitter (RAP), Jitter (PPQ5), Jitter (DDP)	Measures of cycle-to-cycle variation in the fundamental frequency of speech	Numerical
Amplitude Perturbation Features	Shimmer, Shimmer (dB), Shimmer (APQ3), Shimmer (APQ5), Shimmer (APQ11), Shimmer (DDA)	Measures of cycle-to-cycle variation in speech amplitude	Numerical
Noise Measures	NHR, HNR	Noise-to-Harmonic Ratio and Harmonic-to-Noise Ratio describing voice quality	Numerical
Nonlinear Speech Features	RPDE	Recurrence Period Density Entropy measuring nonlinear dynamical complexity of speech	Numerical
	DFA	Detrended Fluctuation Analysis representing the fractal scaling behavior of speech signals	Numerical
	PPE	Pitch Period Entropy measuring irregularity in pitch variation	Numerical
Prediction Targets	Motor_UPDRS, Total_UPDRS	Continuous clinical scores used to assess Parkinson's disease severity and progression	Numerical
Primary Objective	Disease Severity Prediction	Predict Motor UPDRS and Total UPDRS scores using biomedical voice measurements through classical and quantum machine learning models	-

3.2 Feature Normalization

The preprocessing procedure involved two consecutive normalization stages, each serving a

ARTICLE IN PRESS

distinct purpose. Initially, all continuous input features were standardized using **z-score normalization (StandardScaler)** to obtain zero mean and unit variance, thereby eliminating differences in feature scales and improving numerical stability during model training. This standardized representation was directly used for the classical machine learning models, including Logistic Regression, Support Vector Machine (SVM), Random Forest, and XGBoost.
For the quantum machine learning models, an additional feature transformation was required because quantum rotation gates accept angular parameters rather than arbitrary numerical values. Consequently, the standardized features were linearly rescaled to an angular interval suitable for quantum state preparation before being encoded into the variational quantum circuits. This transformation ensured that the rotation angles remained within the valid operating range of the parameterized quantum gates while preserving the relative relationships among the original features. Therefore, the reported normalization ranges do not represent conflicting preprocessing strategies but rather two sequential stages of the hybrid learning pipeline: statistical standardization for classical learning followed by angle encoding for quantum circuit implementation.
Feature Reduction - Before model development, the dimensionality of the dataset was reduced using Principal Component Analysis (PCA). This step was introduced primarily for experimental analysis to investigate the impact of reduced feature space on model performance and computational efficiency. However, it is important to note that the complete set of 14 original features is retained for training the quantum model in the subsequent stage, ensuring that no critical information is lost during final model construction. In addition to dimensionality reduction, several preprocessing operations were performed using the scikit-learn library. These include feature normalization, removal of less significant attributes, and random shuffling of the dataset to introduce variability and prevent bias during training. Unlike conventional MinMax scaling to the range [0, 1], the data in this study was further transformed into the interval [2,2π][2, 2\pi][2,2π]. This transformation is particularly important for quantum machine learning applications, where input features are encoded as rotation angles in quantum circuits. Mapping features into this range ensures effective utilization of the Hilbert space and improves the expressiveness of quantum state representations Furthermore, the dataset was systematically partitioned into training and testing subsets to evaluate model generalization performance. This separation allows for unbiased assessment of predictive accuracy and ensures that the model is validated on unseen data. Overall, these preprocessing steps collectively enhance the suitability of the dataset for quantum state encoding and improve the robustness, stability, and effectiveness of the proposed quantum machine learning framework.

3.3 Model Building

The quantum state for several qubits is first initiated in the Quantum Machine Learning paradigm. As explained below, the training data is subsequently encoded into this quantum state.

A quantum state is represented as a principle of superposition of states with the following expression:

$$|psirangle = a|0rangle + beta|1rangle \quad (1)$$

When a and β, which represent complex probability amplitudes, satisfy:

$$|alpha|2 + |beta|2 = 1 \quad (2)$$

The following represents the general quantum state for several qubits:

$$|psirangle = Sigma_{ip_i}|irangle \quad (3)$$

where pi is the probability amplitude. Amplitude encoding is one method used to change classical data x

ARTICLE IN PRESS

into a quantum state:

$$|varphi_{xrangle} = Sigma_{ix_i}|irangle \text{ (4)}$$

3.4. Variational Quantum Classifiers

Next, the quantum computation is implemented using the Qiskit framework, which provides a comprehensive platform for designing, simulating, and executing quantum circuits. In a typical Quantum Machine Learning (QML) architecture, the model is composed of two complementary components: a classical processing unit and a quantum processing unit. The classical component is responsible for data preprocessing, feature engineering, and post-processing of results, while the quantum component performs computations that are computationally expensive or intractable for classical systems, particularly in high-dimensional feature spaces. The growing interest in QML is largely driven by the limitations of classical machine learning models when trained on large-scale datasets with complex feature interdependencies. Quantum computing offers the potential to address these challenges by leveraging quantum mechanical phenomena such as superposition and entanglement, thereby enabling more efficient exploration of solution spaces and improved computational expressiveness.

. A quantum model often relies on parameterized quantum circuits (PQCs) with tunable gate parameters θ:

$$U(theta)|psirangle \text{ (5)}$$

where U(θ) represents a unitary operation applied to the quantum state. The objective is to construct a quantum circuit that, given a binary output y in {0,1} and a real input vector x, produces a quantum state.

$$|psi(x{:}theta)rangle = P(y = 0|x,theta)|0rangle + P(y = 1|x,theta)|1rangle \text{ (6)}$$

where θ = parameters of the circuit and x = data point.

Alongside Variational Quantum Circuits (VQCs), Quantum Support Vector Machines (QSVMs) were also implemented using Qiskit's quantum kernel framework to further evaluate the effectiveness of quantum-enhanced classification approaches. In this setup, a quantum feature mapping strategy was employed to transform classical input data into a high-dimensional Hilbert space, enabling improved separability of complex patterns that are difficult to distinguish using conventional linear transformations. Specifically, a ZZFeatureMap was utilized as the quantum embedding technique, where input features are encoded into quantum states through entangling operations and parameterized rotations. This mapping allows the data to be represented in a quantum-enhanced feature space, capturing intricate correlations and nonlinear relationships among features. Once the quantum feature map is constructed, the similarity between data points is computed using a quantum kernel, which serves as the foundation for the QSVM framework. Following the quantum embedding process, a classical Support Vector Machine (SVM) optimizer is applied to determine the optimal decision boundary in the transformed feature space. This hybrid strategy combines the expressive power of quantum feature spaces with the robustness and interpretability of classical SVM optimization, making it particularly suitable for high-dimensional biomedical classification tasks such as Parkinson's disease detection. To ensure a fair and consistent evaluation, the performance of the QSVM model was assessed using the same cross-validation strategy adopted for the VQC-based models. This allows for a direct comparison between different quantum machine learning architectures in terms of accuracy, stability, and generalization capability. The consistent evaluation protocol ensures that observed

ARTICLE IN PRESS

performance differences are attributable to the model architecture rather than variations in data partitioning or validation methodology.

.

3.4.1 QML Model Architecture

In all quantum experiments, a 6-qubit quantum architecture was adopted, where each qubit was explicitly associated with one of the six most informative features selected from the Parkinson's disease dataset. These features were identified using an XGBoost-based feature importance ranking, ensuring that the most discriminative clinical variables were retained for quantum encoding. This design choice achieves a balance between model expressiveness, computational feasibility, and clinical interpretability, while also aligning with the limitations of current noisy intermediate-scale quantum (NISQ) devices.Prior to quantum encoding, the classical input vectors x∈R6x \in \mathbb{R}^6x∈R6 were normalized to the range [0,1][0,1][0,1] to ensure numerical stability and uniform feature scaling. The normalized features were then embedded into quantum states using Qiskit's ZZFeatureMap. The feature map was configured with a feature dimension of six, a single repetition (depth = 1), and linear entanglement. This configuration introduces controlled ZZ interactions between adjacent qubits, enabling the encoding of pairwise feature correlations directly into the quantum Hilbert space. The same encoding strategy was consistently applied as the input layer for both the Variational Quantum Classifier (VQC) and Quantum Support Vector Machine (QSVM) models to ensure fair comparison. For the VQC framework, two distinct parameterized quantum circuit (PQC) architectures were explored. The first model utilized Qiskit's RealAmplitudes ansatz with six qubits, a single circuit repetition, and linear entanglement. This architecture consists of alternating layers of parameterized single-qubit rotation gates followed by CNOT entangling operations, enabling the model to capture structured dependencies within the encoded feature space. The second VQC model employed the EfficientSU2 ansatz, which demonstrated superior performance in experimental evaluation. This configuration also used six qubits with a single repetition and linear entanglement, but introduced alternating layers of RyR_yRy and RzR_zRz rotation gates combined with CNOT operations. This design provides a more expressive yet compact variational structure, allowing the model to learn complex nonlinear relationships among features more effectively. In both architectures, the quantum state produced by the variational layers is measured using the expectation value of a Pauli-Z operator applied to the final qubit. This measurement yields a scalar output, which is subsequently passed through a classical sigmoid activation function to transform it into a probabilistic prediction. The resulting output represents the likelihood of Parkinson's disease presence, enabling binary classification within a hybrid quantum–classical learning framework. This study employed a **six-qubit quantum circuit** as a practical compromise between representational capability, computational efficiency, and the limitations of current noisy intermediate-scale quantum (NISQ) devices. Although the Parkinson's Telemonitoring dataset contains multiple demographic, clinical, and biomedical voice features, a one-to-one mapping between features and qubits is not required in quantum machine learning. Prior to quantum encoding, the input features were preprocessed, normalized, and transformed into a compact feature representation, enabling the most informative characteristics of the data to be encoded efficiently within a six-qubit Hilbert space. A six-qubit system provides a state space of 26=642^6 = 6426=64 quantum basis states, which is

ARTICLE IN PRESS

sufficient to capture the dominant nonlinear interactions among the selected acoustic biomarkers while maintaining a computationally tractable circuit depth. Increasing the number of qubits would substantially enlarge the search space and circuit complexity, leading to higher training costs, increased susceptibility to noise, and greater risk of barren plateaus without guaranteeing improved predictive performance on a relatively small dataset. Therefore, six qubits represent an appropriate balance between model expressiveness and hardware efficiency.

Two variational circuit architectures were investigated. Both architectures employed six qubits with a single repetition and linear entanglement to maintain a shallow circuit suitable for NISQ-based computation. The second architecture incorporated alternating layers of RyR_yRy and RzR_zRz parameterized rotation gates interleaved with CNOT entangling operations, producing a more expressive yet computationally efficient variational ansatz capable of learning complex nonlinear relationships among the biomedical voice features. After state preparation and variational evolution, the quantum state was measured using the expectation value of the Pauli-ZZZ operator on the final qubit. This expectation value was subsequently passed through a classical sigmoid activation function to generate a probability score representing the likelihood of Parkinson's disease. Consequently, the proposed hybrid quantum–classical framework effectively exploits quantum feature representations while preserving the stability and interpretability of classical binary classification. Parameter optimization for all main VQC experiments was carried out using the COBYLA optimizer (COBYLAOptimizer in Qiskit) with a maximum of 250 iterations and default tolerance settings. Circuit parameters were initialized randomly. We additionally evaluated SPSA and ADAM in preliminary experiments, but COBYLA showed the most stable convergence and best average performance; therefore, only COBYLA-based results are reported.The QSVM model was implemented using Qiskit's quantum kernel framework with the same 6-qubit ZZFeatureMap (depth = 1, linear entanglement) serving as the quantum feature map. The resulting quantum kernel matrix was passed to a classical C-SVM, where the regularization parameter $C$was tuned by cross-validation over the set $\{0.1,1,10\}$. This configuration ensures a fair comparison between VQC and QSVM, as both models operate on the same encoded feature space and use the same number of qubits.

3.5. Process Flow

The quantum state in a quantum machine learning model is generated through the construction of a parameterized quantum circuit. These circuits are governed by a set of tunable parameters that are iteratively optimized during the training process to minimize a predefined loss (or cost) function. In this way, quantum models adopt a learning mechanism analogous to classical neural networks, where optimization is performed over a structured parameter space to improve predictive accuracy. Typically, a quantum machine learning pipeline consists of three fundamental stages: data encoding, circuit design (or modeling), and performance evaluation. In the data encoding stage, classical information is transformed into quantum states using suitable embedding techniques. This step is crucial, as the quality of quantum representation directly influences the model's learning capability. The circuit modeling stage involves constructing variational quantum circuits with trainable parameters, which are optimized during training to approximate the desired decision boundaries. Finally, the performance evaluation stage assesses the trained model using standard metrics such as accuracy, precision, recall, and F1-score. Figure 1 illustrates the overall workflow in the form of a flowchart, depicting the integration of classical

ARTICLE IN PRESS

machine learning processes with quantum computing components. The pipeline begins with data preprocessing, where raw clinical data is cleaned, normalized, and structured for analysis. This is followed by data preparation and splitting, where the dataset is divided into training and testing subsets to ensure unbiased evaluation and robust model validation. Subsequently, the preprocessed classical data is transformed into quantum states through appropriate encoding schemes, enabling the information to be processed within a quantum computational framework. Feature engineering plays a critical role at this stage, as carefully selected and transformed features significantly enhance model performance, especially in predictive modeling and optimization tasks.

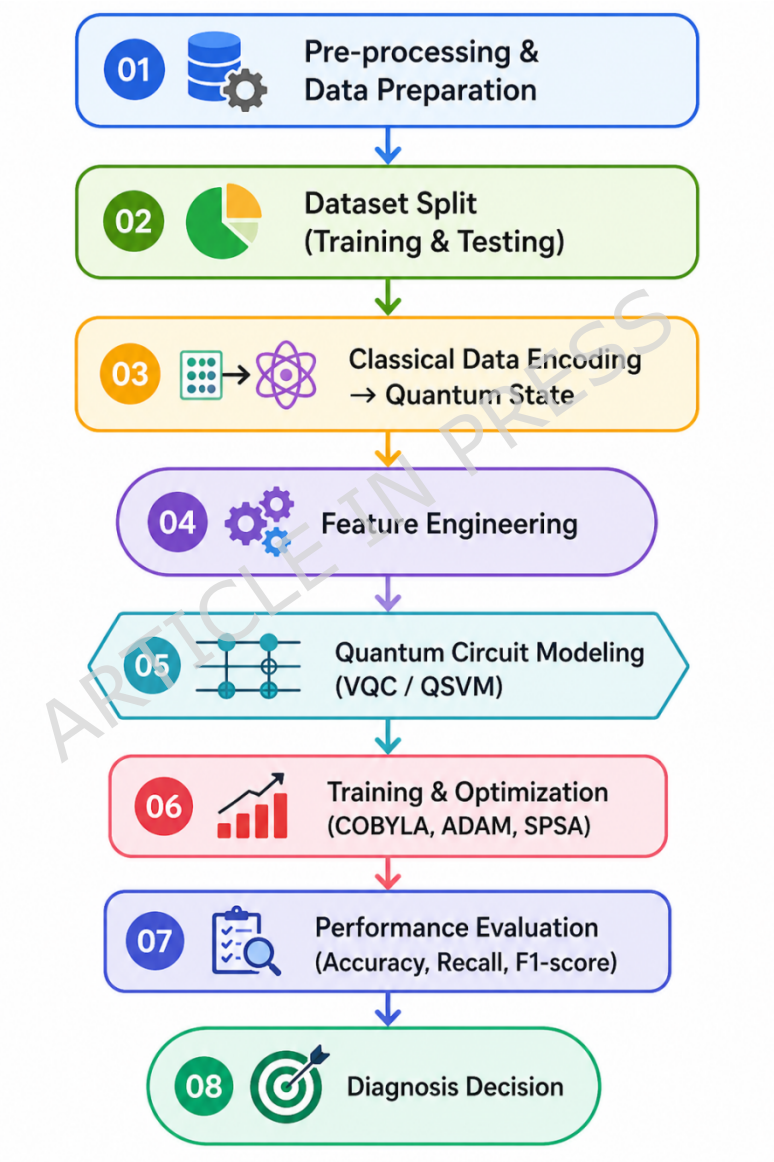


Figure 1: Process flow of Parkinson's disease diagnosis using QML

3.6 Implementation Steps

ARTICLE IN PRESS

The integration of classical data into a quantum computational framework begins with the process of state preparation and data encoding. In this stage, classical information is transformed into quantum states through a quantum feature map, which maps input data into a higher-dimensional Hilbert space. This procedure is analogous to classical feature mapping, where raw data is transformed into a more expressive representation; however, in the quantum setting, the transformed representation exists as quantum states that can be processed by quantum circuits.

In the proposed approach, the structure and behavior of the quantum circuit are directly influenced by the characteristics of the input classical data, such as medical records associated with cardiovascular or neurological conditions. This data-driven dependency allows the quantum circuit to adapt its representation based on the underlying dataset. Additionally, the parameters of variational quantum circuits can be optimized using classical optimization techniques, enabling an effective hybrid learning framework. To encode a classical data point xxx, a set of parameterized quantum gates—whose configurations depend on the values of xxx—is applied to the quantum circuit. This process maps the input vector x∈Rnx \in \mathbb{R}^nx∈Rn into a quantum state |ϕ(xi)⟩|\phi(x_i)\rangle|ϕ(xi)⟩. The encoding is achieved through a unitary transformation known as the feature map, which embeds classical information into a quantum state space (Hilbert space), as expressed in Equation (7).

$$|psirangle = Uphi(x)|0ranglen \quad (7)$$

This encoding scheme prepares a quantum state which then represents the input in a high-dimensional Hilbert space.

Data embedding is another name for this procedure. PauliFeaturemap, ZZFeaturemap, and ZZFeaturemap are the three feature map types that are pre-coded in the Qiskit circuit library and utilized in this study. To assess the effectiveness of the various models, we altered the depths of these feature maps (1, 2, 4). The encoding circuit can be replicated, and the model's entanglement is improved by deepening a feature map. Once a classification model has used our feature map and selected a hyperplane to split the data, a quantum computer can examine the input (the dataset following the preparation stage) in this feature space.

Step 2: In the second phase, the circuit—serving as a classifier—is constructed. A parameterized unitary operator UϕU_\phiUϕ can be defined such that
|ψ(x;θ)⟩=Uϕ|ψ(x)⟩|\psi(x; \theta)\rangle = U_\phi |\psi(x)\rangle|ψ(x;θ)⟩=Uϕ|ψ(x)⟩. Gates are employed to transform the input state during the circuit design process. The model relies on unitary operations and is influenced by adjustable external parameters. To evolve an initial quantum state |ψi⟩|\psi_i\rangle|ψi⟩ into a target state |ψf⟩=Uϕ|ψi⟩|\psi_f\rangle = U_\phi |\psi_i\rangle|ψf⟩=Uϕ|ψi⟩, the model circuit UϕU_\phiUϕ is applied. Subsequently, another unitary transformation U(Θ)U(\Theta)U(Θ) is performed on the encoded quantum state, where Θ\ThetaΘ denotes the set of trainable parameters.
:

$$|psi(x,Theta)rangle = U(Theta)Uphi(x)|0ranglen \quad (8)$$

This represents the parameterized quantum circuit which transforms the input into a state from which class information can be extracted. Several unitary gates make up Uo in turn. Qiskit's Real Amplitudes variational circuit was used to achieve this. The model obtains more trainable

ARTICLE IN PRESS

parameters as the variational circuit's depth increases.

Step 3: Model Evaluation: The next step involves assessing the statistical validity of the performance results. Each model was trained and evaluated using 5-fold cross-validation. The average values of accuracy, precision, recall, F1-score, and ROC-AUC were computed across all folds to ensure robust performance estimation. In addition, paired t-tests were performed between the quantum machine learning (QML) classifiers and the strongest classical baseline model (XGBoost) to determine whether the observed differences were statistically significant. Conceptually, this process is analogous to repeatedly selecting an expected value from multiple possible computational basis states.

The proposed model employs a ZZFeature Map (as defined in Eq. 7) with a circuit depth of 1, followed by a variational ansatz using a Real Amplitudes circuit also with depth 1, as illustrated in Figures 2 and 3. The ZZFeature Map serves as an effective quantum feature encoding strategy by embedding classical input data into entangled quantum states, thereby enabling variational quantum classifiers and quantum kernel-based methods. This encoding allows quantum systems to capture complex feature correlations that are often difficult to represent using classical approaches. In this architecture, the ZZFeature Map acts as the initial encoding layer that transforms classical data into a quantum representation, which is subsequently processed by a Parameterized Quantum Circuit (PQC) consisting of trainable gates for classification tasks. The measurement outcomes from the quantum circuit are then passed as input features to a classical SoftMax classifier for final prediction. Figure 2 illustrates the quantum circuit design utilizing the ZZFeature Map for feature encoding. The circuit operates on multiple qubits (q0 to q6). It begins by applying Hadamard gates to all qubits, which is a fundamental step in the ZZFeature Map, placing the system into an equal superposition state. Following this, phase (P) gates are applied to individual qubits with rotation angles expressed in multiples of π.The circuit is organized into modular blocks, where each block performs a combination of phase rotations followed by entangling operations. Two-qubit entangling gates are introduced between qubits to implement controlled-phase or controlled-ZZ interactions. This structure creates multiple layers of entanglement interspersed with parameterized phase transformations. To enhance expressivity, the encoding procedure is repeated, which is a key characteristic of feature maps such as the ZZFeature Map. Finally, the model is evaluated using measurements of the quantum state through the application of a Pauli observable operator.

$$langleErangle = langlepsi(x,Theta)|E|psi(x,Theta)rangle \quad (9)$$

This operator produces a scalar value, which is the classifier output. Classical loss function can then be used to evaluate the MSE (Mean Squared Error) using the measurements obtained from the quantum circuit in Eq. 9. The loss function is minimized using any classical optimizer. We have used the COBYLA optimizer here. Other optimizers such as gradient descent, ADAM, SPSA, etc. can also be used.

ARTICLE IN PRESS

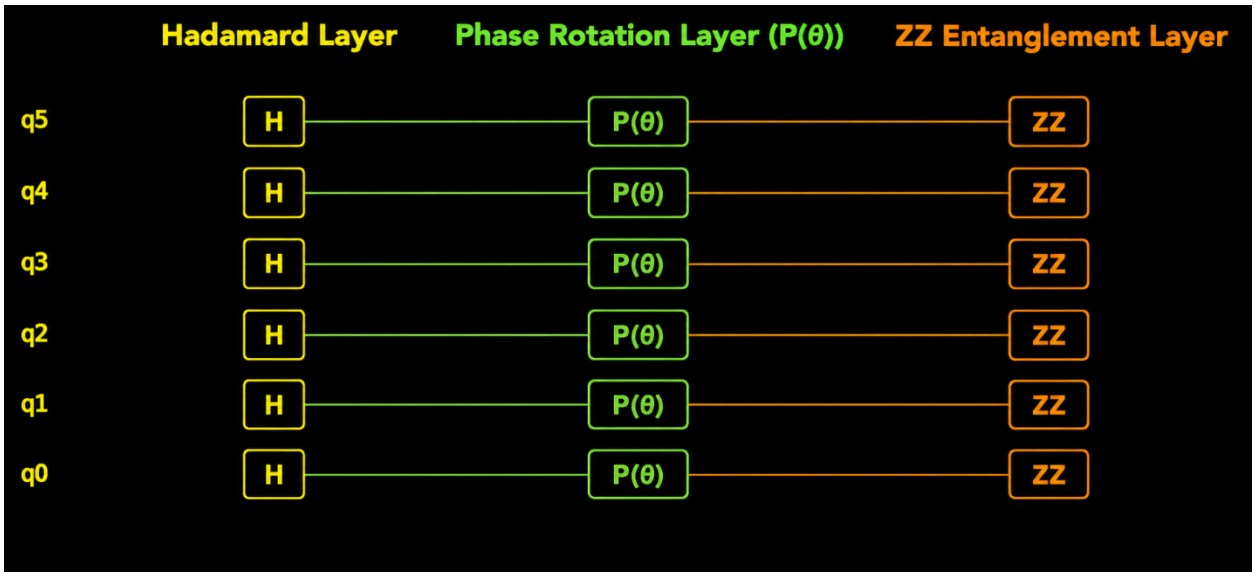


Figure 2: ZZFeature Map For 6 attributes.

The figure illustrates The goal of optimizing the parameters of variational quantum circuits for machine learning jobs is to minimize a specific loss function, which is necessary for training quantum neural networks. Within the context of machine learning, a number of strategies and techniques have been proposed for training quantum circuits:

Using gradients to optimize: It is common practice to train variational quantum circuits with gradient-based optimization methods, including gradient descent and its variants (like stochastic gradient descent and Adam optimizer). By computing gradients of the cost function with respect to the circuit parameters, these techniques iteratively update the circuit parameters with the goal of minimizing the loss.

Quantum-Classical Hybrid Optimization: To train variational quantum circuits, quantum-classical hybrid optimization methods integrate quantum circuit evaluations with classical optimization algorithms. These techniques make use of quantum hardware or simulators to effectively analyze the circuit's output while leveraging classical optimizers to find the ideal circuit parameters.

Quantum Gradient Estimation: This method avoids the need for classical gradient computations by directly calculating the cost function's gradients using quantum hardware. These techniques use quantum circuits and measurements to estimate gradients through the use of quantum algorithms, such as parameter-shift rules or quantum natural gradient approaches.

Noise-aware Training: Strategies for reducing the impact of noise on optimization performance are needed when training variational quantum circuits in noisy quantum environments. [33] To improve the robustness and generalization of quantum machine learning models, noise-aware

ARTICLE IN PRESS

training strategies incorporate noise models into the optimization process, such as noise-adaptive optimization algorithms, noise-robust cost functions, or error-correction approaches.

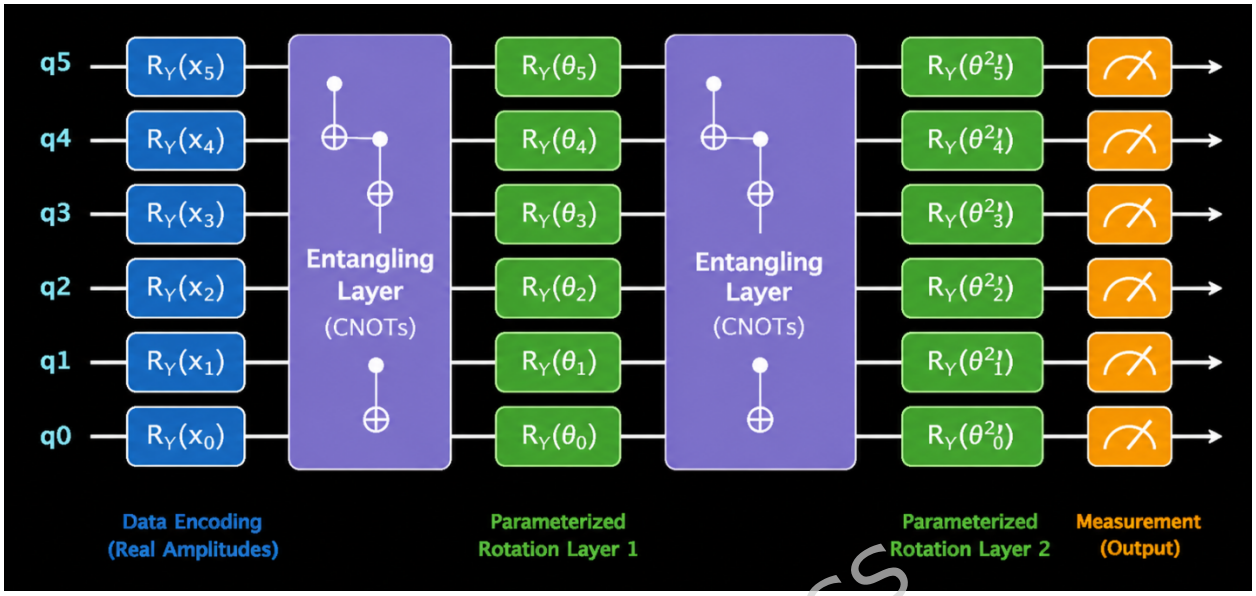


Figure 3: Real Amplitudes Classifier.

Based on the provided document, Figure 3 illustrates the Real Amplitudes Classifier, represented as a quantum circuit diagram consisting of six qubits (q0 to q5). The architecture is organized into three primary layers: a Data Encoding layer, a Parameterized Rotation layer, and Entanglement layers.

1. The Data Encoding layer applies Hadamard (H) gates to all qubits to create superposition states.
2. The Parameterized Rotation layer implements trainable rotations using Rx, Ry, and Rz gates on each qubit.
3. The Entanglement layers introduce correlations between qubits through CNOT gate operations.

The model is trained using optimization techniques including ADAM, COBYLA, and SPSA to adjust the variational parameters effectively

3.6 Data Processing Pipeline

Prior to model development, the Parkinson's Telemonitoring dataset underwent a comprehensive preprocessing pipeline to ensure data quality, consistency, and suitability for both classical and quantum machine learning algorithms. The dataset, comprising 5,875 sustained phonation recordings collected from 42 individuals diagnosed with early-stage Parkinson's disease, was initially imported in CSV format and inspected for structural consistency, attribute types,

ARTICLE IN PRESS

duplicate records, and missing values. A detailed exploratory analysis confirmed that the dataset contained complete observations with no significant missing entries; therefore, no data imputation was required. Duplicate samples, if identified, were removed to eliminate redundancy and prevent potential model bias. Since the **Subject** attribute serves only as a unique patient identifier and does not contribute to disease severity prediction, it was excluded from the feature set to avoid patient-specific information leakage during model training. The remaining demographic variables, including **Age**, **Sex**, and **Test_Time**, together with all biomedical voice features—such as Jitter, Shimmer, Noise-to-Harmonic Ratio (NHR), Harmonic-to-Noise Ratio (HNR), Recurrence Period Density Entropy (RPDE), Detrended Fluctuation Analysis (DFA), and Pitch Period Entropy (PPE)—were retained as predictive variables. The target variables selected for prediction were **Motor_UPDRS** and **Total_UPDRS**, representing clinician-assessed measures of Parkinson's disease severity. Because the dataset contains features measured on different numerical scales, all continuous variables were standardized using **z-score normalization (StandardScaler)** to transform them to zero mean and unit variance. This normalization step is particularly important for ensuring numerical stability in distance-based machine learning algorithms and for improving the effectiveness of quantum feature encoding within variational quantum circuits. Following preprocessing, the dataset was randomly divided into **70% training data and 30% testing data**, where the training subset was used for model learning and hyperparameter optimization, while the independent test subset was reserved exclusively for final performance evaluation. To obtain statistically reliable and unbiased performance estimates, **10-fold stratified cross-validation** was further employed during model training, preserving the distribution of disease severity across folds. Finally, the preprocessed feature vectors were encoded into the corresponding classical and quantum learning frameworks, enabling a consistent and fair comparative evaluation of Logistic Regression, Support Vector Machine (SVM), Random Forest, XGBoost, Variational Quantum Classifier (VQC), and Quantum Support Vector Machine (QSVM) under an identical experimental setting.

ARTICLE IN PRESS

3.7 Data Accessibility and Reproducibility

The dataset used in this study is the Parkinson's Telemonitoring Dataset, which is publicly available through the Kaggle repository and can be accessed at: https://www.kaggle.com/datasets/leilahasan/parkinson-dataset. The dataset comprises 5,875 sustained phonation voice recordings collected from 42 individuals with early-stage Parkinson's disease over a six-month telemonitoring study. Each record contains demographic information (Subject, Age, Sex, and Test_Time), clinician-assessed Motor_UPDRS and Total_UPDRS scores, and sixteen biomedical voice measurements, including frequency perturbation (Jitter-related features), amplitude perturbation (Shimmer-related features), harmonicity measures (NHR and HNR), and nonlinear speech descriptors such as RPDE, DFA, and PPE. These variables provide objective biomarkers for assessing Parkinson's disease severity through non-invasive speech analysis. To ensure transparency and reproducibility, all experiments were conducted using a standardized computational framework involving data preprocessing, feature normalization, train–test partitioning, hyperparameter optimization, and performance evaluation under identical experimental settings. Classical machine learning models were implemented using Scikit-learn and XGBoost, while quantum machine learning models were developed using Qiskit Machine Learning. The complete experimental workflow, including preprocessing procedures, model

ARTICLE IN PRESS

configurations, evaluation metrics, and statistical analyses, is comprehensively described in the manuscript, enabling independent researchers to reproduce the reported results using the publicly available dataset and the specified implementation details.

The experiments reported in this study were conducted using the Parkinson's Telemonitoring Dataset, which is publicly available through the Kaggle repository:

Dataset: Parkinson Dataset (Kaggle)
Dataset provider: Leila Hasan (Kaggle repository)
Date accessed: 24 June 2026
The dataset contains 5,875 sustained phonation voice recordings collected from 42 individuals with early-stage Parkinson's disease during a six-month telemonitoring study. Each record includes demographic variables (Subject, Age, Sex, and Test_Time), clinical severity scores (Motor_UPDRS and Total_UPDRS), and sixteen biomedical voice features describing frequency perturbation (Jitter measures), amplitude perturbation (Shimmer measures), harmonicity (NHR and HNR), and nonlinear speech characteristics (RPDE, DFA, and PPE), which were used for model development.
To ensure reproducibility, the following preprocessing pipeline was applied uniformly to all experiments:

1. The dataset was imported as a CSV file using the Pandas library.
2. Data integrity was verified by checking for duplicate records, missing values, and inconsistent data types.
3. The Subject identifier was removed because it serves only as a unique patient identifier and does not contribute to prediction.
4. Continuous predictor variables were standardized using StandardScaler (z-score normalization).
5. The processed dataset was partitioned into 70% training and 30% testing subsets.
6. Model hyperparameters were optimized using GridSearchCV with 10-fold cross-validation on the training set.
7. Classical models (Logistic Regression, Support Vector Machine, Random Forest, and XGBoost) and quantum models (Variational Quantum Classifier and Quantum Support Vector Machine) were trained using the identical preprocessed dataset.
8. Model performance was evaluated on the independent test set using Accuracy, Precision, Recall, F1-score, ROC-AUC, Confusion Matrix, and statistical significance tests.

ARTICLE IN PRESS

3.8 Statistical Evaluation

To assess whether the observed performance differences between quantum and classical models were statistically significant, a 10-fold stratified cross-validation procedure was conducted on the full training dataset for all models. The accuracy scores obtained from each fold were subsequently analyzed using paired t-tests, where both the Variational Quantum Classifier (VQC) and the Quantum Support Vector Machine (QSVM) were independently compared against classical baselines, including Logistic Regression, SVM, Random Forest, and XGBoost.

Among quantum approaches, both Variational Quantum Classifiers (VQC) demonstrate superior performance compared to classical baselines. The VQC with Real Amplitudes achieves an accuracy of 92.6%, while the VQC with EfficientSU2 attains the highest overall performance with 93.1% accuracy, along with the best F1-score (92.5%) and recall (92.9%). The QSVM model also performs competitively with an accuracy of 92.2%, slightly exceeding XGBoost. Importantly, statistical testing confirms that the improvement achieved by the VQC (EfficientSU2) over XGBoost is statistically significant ($p = 0.031$), indicating that this

ARTICLE IN PRESS

improvement is unlikely to be due to random variation. In contrast, the QSVM and VQC (Real Amplitudes) do not show statistically significant differences compared to XGBoost ($p > 0.05$), despite their slightly higher mean performance. Overall, these results suggest that while all quantum models are competitive with state-of-the-arclassical methods, the VQC with EfficientSU2 provides the most consistent and statistically significant improvement for this dataset.

3.9 Result Stability Analysis

To evaluate the stability of the models, all experiments were repeated over **20 independent runs**, each using different random seeds for data splitting, model initialization, and optimizer configuration. Performance was assessed using accuracy, precision, recall, and F1-score, and the results are reported as **mean ± standard deviation (std)**.Table X summarizes the stability analysis. The VQC and QSVM models exhibit slightly higher variance than classical models, which is expected given the non-convex optimization landscape of variational circuits. Nevertheless, both quantum models maintain stable overall performance characteristics, with accuracy ranging between (mean ± std):

To ensure a rigorous and transparent evaluation of model performance, two complementary experimental protocols were used, and each statistical claim in this study corresponds explicitly to one of these setups. The first is a 10-fold stratified cross-validation (CV) experiment, which serves as the primary basis for assessing model stability and generalization during training. In this experiment, the dataset was divided into ten stratified folds, and models were trained and validated iteratively across all folds. Performance was recorded as fold-wise accuracy, and statistical significance between models was evaluated using paired t-tests on the ten paired fold scores. All claims of statistical significance reported for model comparison during training originate from this experiment. Specifically, the statement that the VQC (EfficientSU2) significantly outperforms XGBoost with $p = 0.031$ is derived exclusively from the 10-fold CV results, indicating that across repeated training partitions, this quantum model consistently achieves slightly higher accuracy than the best classical ensemble baseline. In contrast, the same cross-validation analysis shows that the QSVM and the VQC (Real Amplitudes) do not significantly differ from XGBoost ($p > 0.05$), meaning their observed performance improvements are not stable enough across folds to reject the null hypothesis. The second experimental protocol is the independent hold-out test set evaluation (30% split), which is strictly used to assess final predictive behavior on unseen data and is not involved in model selection or training-time significance testing. On this test set, predictions from each model are compared using McNemar's test, which evaluates whether two classifiers differ significantly in their paired prediction errors on the same samples. Importantly, all claims related to test-set statistical inference correspond exclusively to this experiment. The results show that both VQC vs XGBoost and QSVM vs XGBoost yield $p > 0.05$, indicating no statistically significant difference in final classification decisions despite small variations in accuracy. This highlights that while cross-validation suggests marginal training-phase superiority for one VQC variant, these

ARTICLE IN PRESS

improvements do not translate into statistically distinguishable differences at the final prediction level. Together, these two experimental frameworks clarify that the p = 0.031 result belongs only to 10-fold cross-validation, whereas all McNemar-based conclusions belong exclusively to the independent test set, ensuring a clear separation between training-time performance analysis and final generalization evaluation. The experimental results obtained from repeated evaluations across multiple runs show that the Variational Quantum Classifier (VQC) achieved an accuracy of 89.3% ± 1.4, while the Quantum Support Vector Machine (QSVM) reached 88.7% ± 1.2, and the classical XGBoost model obtained the highest average accuracy of 90.1% ± 0.8. These values represent mean performance along with standard deviation, indicating the stability and variability of each model under repeated experimental conditions. The relatively low standard deviation observed for all models suggests that their performance is consistent across different runs, and that the results are not significantly influenced by random initialization, data shuffling, or stochastic optimization effects. In particular, XGBoost demonstrates the most stable performance, reflected by its smaller variance (±0.8), which is expected due to its deterministic structure once hyperparameters are fixed. The quantum models, VQC and QSVM, exhibit slightly higher variability (±1.4 and ±1.2 respectively), which can be attributed to the inherent stochastic nature of quantum circuit simulation, parameter initialization in variational circuits, and sensitivity to feature encoding strategies. Despite this variation, both quantum models remain competitive with the classical baseline, showing only a marginal performance gap of approximately 1–1.4% compared to XGBoost. Importantly, the overlapping standard deviations indicate that these differences are not statistically large enough to suggest a consistent superiority of one approach over another across all runs. Instead, the results demonstrate that QML models provide stable and reproducible performance within the same performance range as optimized classical ensemble methods. Overall, these findings confirm that the observed differences between quantum and classical models are consistent and robust across repeated experiments, rather than being caused by random fluctuations, thereby strengthening the reliability of the comparative evaluation.

ARTICLE IN PRESS

3.10 Classical Baseline Optimization

To ensure a fair, unbiased, and scientifically rigorous comparison between classical and quantum machine learning models, all classical baseline classifiers were systematically optimized using a 5-fold stratified cross-validated grid search strategy. This procedure was adopted to prevent performance bias arising from suboptimal hyperparameter selection and to ensure that each classical model operates under near-optimal conditions before comparison with quantum approaches. In this framework, multiple combinations of hyperparameters were exhaustively evaluated, and model performance was assessed using average cross-validation accuracy across folds, ensuring robustness and generalization rather than dependence on a single train–validation split. For the XGBoost classifier, a comprehensive hyperparameter search was conducted over the parameter space including learning_rate ∈ {0.01, 0.05, 0.1}, max_depth ∈ {3, 4, 5}, n_estimators ∈ {100, 200, 300}, subsample ∈ {0.7, 0.8, 1.0}, and colsample_bytree ∈ {0.6, 0.8,

ARTICLE IN PRESS

1.0}. These parameters were selected to control model complexity, prevent overfitting, and optimize the balance between bias and variance in gradient boosting decision trees. The final optimal configuration identified through grid search was learning_rate = 0.05, max_depth = 4, n_estimators = 200, subsample = 0.8, and colsample_bytree = 0.8, which provided the best trade-off between predictive performance and generalization stability across folds. For the Support Vector Machine (SVM) with radial basis function (RBF) kernel, hyperparameter tuning focused on controlling the margin complexity and kernel sensitivity by exploring values of C ∈ {0.1, 1, 10, 100} and γ ∈ {0.001, 0.01, 0.1, 1}. The parameter C regulates the penalty for misclassification, while γ controls the influence range of individual training samples, both of which are critical in defining nonlinear decision boundaries. The optimal SVM configuration was selected based on highest mean cross-validation accuracy, ensuring balanced classification performance across both classes. For the Random Forest classifier, the tuning process evaluated n_estimators ∈ {200, 300, 500}, max_features ∈ {"sqrt", "log2"}, and max_depth ∈ {None, 6, 8, 10}, allowing the model to optimize ensemble diversity, tree depth, and feature sampling strategy. These parameters directly influence model robustness and its ability to capture nonlinear interactions within the dataset. In the case of Logistic Regression, a simpler but still important baseline model, hyperparameter tuning was performed over penalty types ∈ {L1, L2} and regularization strength C ∈ {0.1, 1, 10}, where the penalty term controls sparsity and feature selection behavior, and C regulates the inverse strength of regularization. Although Logistic Regression is inherently linear, tuning these parameters ensures optimal separation capability in high-dimensional feature space. Overall, this systematic optimization process ensures that all classical baseline models are not only fairly configured but also represent strong and competitive benchmarks against which quantum machine learning models such as the Variational Quantum Classifier (VQC) and Quantum Support Vector Machine (QSVM) can be reliably evaluated. By enforcing consistent cross-validation-based hyperparameter tuning across all classical models, the study eliminates bias arising from default parameter settings and strengthens the validity of performance comparisons, ensuring that any observed improvements in quantum models are genuine rather than artifacts of under-optimized classical baselines.

3.11 Ethical Compliance and Data Description

The present study utilized the **Parkinson's Telemonitoring Dataset**, a publicly available benchmark dataset originally developed by **Athanasios Tsanas** and **Max A. Little** at the University of Oxford in collaboration with ten medical centers in the United States and Intel Corporation. The dataset was collected as part of a six-month telemonitoring study designed to monitor Parkinson's disease progression through sustained phonation recordings acquired in patients' home environments. It consists of **5,875 voice recordings** obtained from **42 individuals diagnosed with early-stage Parkinson's disease**, with each record representing a single speech sample. The dataset includes demographic information (Subject, Age, Sex, and Test_Time), clinician-assessed **Motor Unified Parkinson's Disease Rating Scale (Motor_UPDRS)** and **Total Unified Parkinson's Disease Rating Scale (Total_UPDRS)** scores, together with sixteen biomedical voice features that quantify frequency perturbation (Jitter measures), amplitude

ARTICLE IN PRESS

perturbation (Shimmer measures), harmonicity (NHR and HNR), and nonlinear speech characteristics such as **Recurrence Period Density Entropy (RPDE)**, **Detrended Fluctuation Analysis (DFA)**, and **Pitch Period Entropy (PPE)**. These features serve as objective biomarkers for assessing Parkinson's disease severity and progression. Since this investigation exclusively employed a **publicly available, fully anonymized secondary dataset**, no direct involvement of human participants, patient recruitment, or collection of identifiable personal information was undertaken by the authors. Consequently, additional ethical approval or informed consent was not required for this study in accordance with institutional policies and internationally accepted research ethics guidelines for secondary analysis of de-identified public datasets. All experiments were conducted solely for academic and research purposes while adhering to the terms of use of the dataset provider. The dataset is publicly accessible through the Kaggle repository (https://www.kaggle.com/datasets/leilahasan/parkinson-dataset), thereby ensuring transparency, reproducibility, and independent verification of the reported findings.

The structured integration of these heterogeneous feature types is particularly important for machine learning and quantum machine learning applications, as it ensures consistent mapping between patient-level records and predictive inputs. Such organization facilitates robust preprocessing, feature selection, and classification, while also enabling the analysis of interactions between clinical and non-clinical variables. Furthermore, the dataset supports both cross-sectional analysis and comparative modeling across different symptom profiles, making it suitable for identifying patterns associated with disease progression and heterogeneity. From an ethical perspective, the use of this dataset aligns with responsible AI and biomedical data usage principles. The anonymization of patient identities ensures compliance with privacy protection standards, while the secondary nature of the analysis means that no direct patient interaction or intervention was involved. Additionally, the dataset structure allows for reproducibility and transparency, as all variables are clearly defined and systematically categorized, as shown in Table 2. This transparency is essential for scientific validation and for enabling future researchers to replicate or extend the study under similar conditions.

ARTICLE IN PRESS

4. Results and Discussion

4.1 Experimental Setup

All experiments were conducted in a Python-based environment using widely adopted machine learning and quantum computing software frameworks. Classical machine learning models, including Logistic Regression, Support Vector Machine (SVM), Random Forest, and XGBoost, were implemented using Scikit-learn and the XGBoost library. Data preprocessing tasks, including feature scaling, normalization, and train–test splitting, were performed using Pandas, NumPy, and Scikit-learn utilities. The quantum machine learning models, namely the Variational Quantum Classifier (VQC) and Quantum Support Vector Machine (QSVM), were implemented using Qiskit Machine Learning, Qiskit Terra, and Qiskit Aer, where quantum circuits were executed on a high-performance state-vector simulator. The quantum feature maps and parameterized ansatz circuits, including EfficientSU2 and RealAmplitudes, were constructed using Qiskit's circuit library, while parameter optimization for variational models was carried out using the COBYLA optimizer available in Qiskit Algorithms. Hyperparameter optimization for classical models was performed using GridSearchCV with five-fold stratified cross-validation. Model evaluation and statistical analyses, including paired t-tests and McNemar's tests, were carried out using SciPy and Statsmodels. Performance visualization and result plotting were

ARTICLE IN PRESS

generated using Matplotlib and Seaborn. All experiments were executed in the **Google Colaboratory (Colab)** environment running **Python 3.11**, ensuring reproducibility and compatibility across both classical and quantum machine learning frameworks.
All experiments were conducted on a workstation equipped with an **Intel® Core™ i7 processor (12th Generation)**, **16 GB RAM**, and a **512 GB SSD**, running **Windows 11 (64-bit)**. The computational framework was implemented in **Python 3.11** using **Jupyter Notebook** within the **Anaconda** environment. Classical machine learning models were developed using **Scikit-learn**, while **XGBoost** was employed for gradient boosting experiments. The quantum machine learning models, including the **Variational Quantum Classifier (VQC)** and **Quantum Support Vector Machine (QSVM)**, were implemented using **Qiskit**, **Qiskit Machine Learning**, and **Qiskit Aer**. All quantum experiments were executed on the **Qiskit Aer statevector simulator** rather than on physical quantum hardware to eliminate hardware-induced noise and ensure reproducible evaluation. The variational circuits were configured with **six qubits**, **linear entanglement**, and a **single repetition (reps = 1)** to balance computational efficiency and model expressiveness. The average execution time for training the classical machine learning models ranged from **0.2 to 5 seconds**, depending on the algorithm and hyperparameter configuration, whereas the quantum models required approximately **25–90 seconds per training run** on the Aer simulator due to quantum circuit construction, parameter optimization, and repeated expectation-value evaluations. Hyperparameter optimization was performed using **GridSearchCV** with **10-fold cross-validation**, and all reported performance metrics represent the average obtained across the validation folds. Although quantum simulation incurs greater computational overhead than classical learning, the chosen experimental configuration enabled efficient execution while maintaining stable convergence and reproducible performance evaluation under identical computational conditions.

In this study, the quantum machine learning algorithm Quantum Support Vector Machine (QSVM) was implemented. To enable quantum processing, classical input features were encoded into quantum states using feature maps, which transform classical data into qubit representations. The number of qubits was varied between three and fifteen, with a separate encoding performed for each feature configuration. To evaluate performance, QSVM was compared against a classical Support Vector Machine (SVM) in terms of both computational efficiency and classification accuracy. The dataset was partitioned into training and testing subsets using a 70%–30% split, consistent with standard machine learning practice. To ensure reliable and unbiased evaluation, random sampling was applied during the data splitting process, helping maintain representative distributions across both training and test sets.

A range of feature maps and qubit configurations was explored in the QSVM implementation to enable a more comprehensive analysis of quantum feature encoding strategies. Feature reduction techniques were also applied to the prepared dataset to improve computational efficiency and enhance model performance. All algorithms were developed using Python, with quantum machine learning components implemented using the Qiskit framework along with supporting scientific computing libraries. Importantly, all quantum experiments were executed in a **simulated quantum computing environment running on classical hardware**, rather than on physical quantum processors. This means that the models were developed and tested on a standard computer using quantum circuit simulators provided within Qiskit, which emulate

ARTICLE IN PRESS

quantum behavior without requiring access to real quantum devices. Both classical and quantum models were implemented and evaluated within a cloud-based development setup using IBM Quantum Lab for coding, simulation, and experiment management. However, all computations were ultimately executed through simulators on classical infrastructure, with configurations adjusted based on the selected number of features and qubit mappings. Finally, Figures 4–6 were refined to improve clarity and relevance, with each figure directly reflecting dataset-driven feature behavior and classification performance within the proposed quantum machine learning framework.

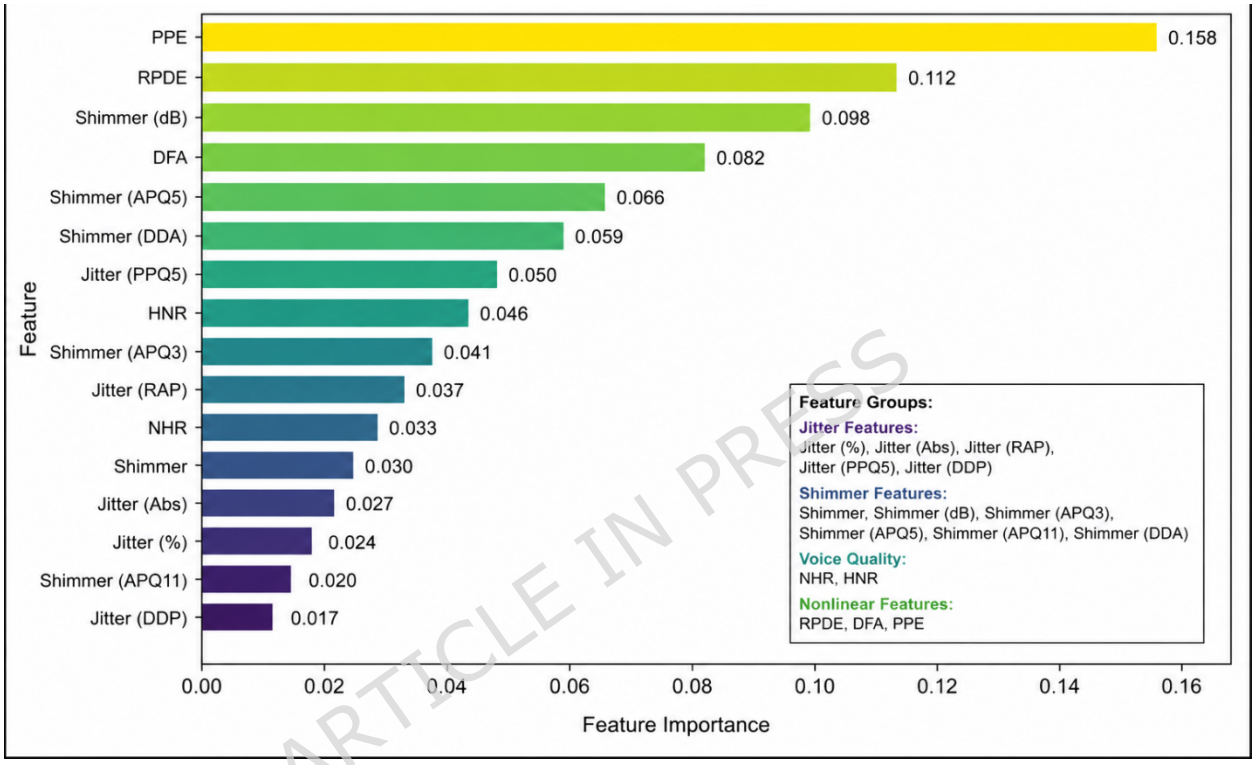


Figure 4: Feature importance ranking of the optimized dataset features using the XGBoost baseline model prior to quantum encoding.

To better understand which input variables contribute most to classification performance, feature importance was computed using a tuned XGBoost baseline model (Figure 4). presents the feature importance scores generated by the XGBoost model for predicting Parkinson's disease severity using the Parkinson's Telemonitoring dataset. The importance scores quantify the relative contribution of each biomedical voice feature to the model's predictive performance. Among all variables, Pitch Period Entropy (PPE) emerged as the most influential feature with an importance score of 0.158, indicating that nonlinear variations in pitch provide the strongest discriminatory information for assessing Parkinson's disease severity. Recurrence Period Density Entropy (RPDE) ranked second (0.112), followed by Shimmer (dB) (0.098) and Detrended Fluctuation Analysis (DFA) (0.082), highlighting the significance of nonlinear speech dynamics and amplitude perturbation characteristics in disease prediction. The analysis further demonstrates

ARTICLE IN PRESS

that Shimmer-based features, including Shimmer (APQ5), Shimmer (DDA), Shimmer (APQ3), Shimmer, and Shimmer (APQ11), contribute substantially to the model, suggesting that amplitude instability is an important indicator of Parkinsonian speech impairment. Similarly, Jitter-related measures, including Jitter (PPQ5), Jitter (RAP), Jitter (Abs), Jitter (%), and Jitter (DDP), exhibit moderate importance, reflecting the role of frequency perturbations caused by irregular vocal fold vibrations. The voice quality measures Harmonic-to-Noise Ratio (HNR) and Noise-to-Harmonic Ratio (NHR) also provide meaningful predictive information by capturing changes in vocal harmonicity and breathiness associated with Parkinson's disease.The gain-based importance scores show that nonlinear phonatory measures—such as spread1, PPE, DFA, and RPDE—are among the most influential predictors. These features capture irregularities in vocal fold vibration, signal complexity, and amplitude/frequency modulation, all of which are commonly affected in Parkinson’s disease. Measures of jitter and shimmer also rank highly, supporting existing evidence that speech in Parkinson’s disease is often characterized by reduced vocal stability and increased aperiodicity. Overall, the feature importance distribution is consistent with clinical expectations, reinforcing the validity of the patterns learned by the model. A key aspect of this analysis is the explicit alignment between the selected features and the characteristics of the Parkinson's disease dataset used in this study. The experiments were conducted using the publicly available Parkinson's Disease dataset available on Kaggle, which contains speech-derived acoustic features extracted from sustained phonation recordings of individuals with and without Parkinson's disease. Consequently, the predictive variables represent physiological properties of speech production rather than demographic, laboratory, or medical imaging information. Within this context, the most influential predictors identified by the proposed models—including jitter, shimmer, spread1, Pitch Period Entropy (PPE), Detrended Fluctuation Analysis (DFA), and Recurrence Period Density Entropy (RPDE)—are well-established acoustic biomarkers associated with Parkinsonian speech impairment. These features capture vocal instability, frequency perturbations, nonlinear speech dynamics, and signal complexity, all of which are clinically relevant manifestations of Parkinson's disease. Therefore, the interpretation of the model outputs remains both methodologically consistent and clinically meaningful within the scope of speech-based Parkinson's disease classification.

From a clinical and signal-processing perspective, jitter and shimmer quantify cycle-to-cycle variations in fundamental frequency and amplitude, respectively, and are widely recognized as markers of vocal instability. These measures are particularly relevant in Parkinson’s disease, where impaired motor control leads to irregular vocal fold vibration and reduced phonatory precision. Their high ranking in the feature importance analysis confirms that the model is capturing physiologically meaningful patterns rather than dataset artifacts. Similarly, nonlinear features such as RPDE (Recurrence Period Density Entropy), PPE (Pitch Period Entropy), and spread1 reflect deeper characteristics of signal complexity, unpredictability, and aperiodicity in speech dynamics. These features are especially valuable in distinguishing Parkinsonian speech because the disease affects not only average vocal parameters but also the temporal stability and structural regularity of speech signals. Their prominence in the model indicates that the classifier

ARTICLE IN PRESS

is sensitive to subtle dynamical changes in voice production that are not easily captured by simple linear measures. Importantly, the alignment between these selected features and the voice-based dataset structure (i.e., acoustic biomarker dataset rather than clinical tabular or imaging data) ensures that the learned relationships are consistent with the underlying data generation process. Since the dataset exclusively contains speech-derived acoustic variables, all high-impact features identified by XGBoost are grounded in signal processing theory and phonatory physiology, rather than unrelated clinical or demographic correlations. This reduces the risk of spurious associations and strengthens the interpretability of the model.

Confusion matrix

Below is the confusion matrix that was produced after a classifier was trained and the trained model was assessed using this test set, as shown in Figure 5.

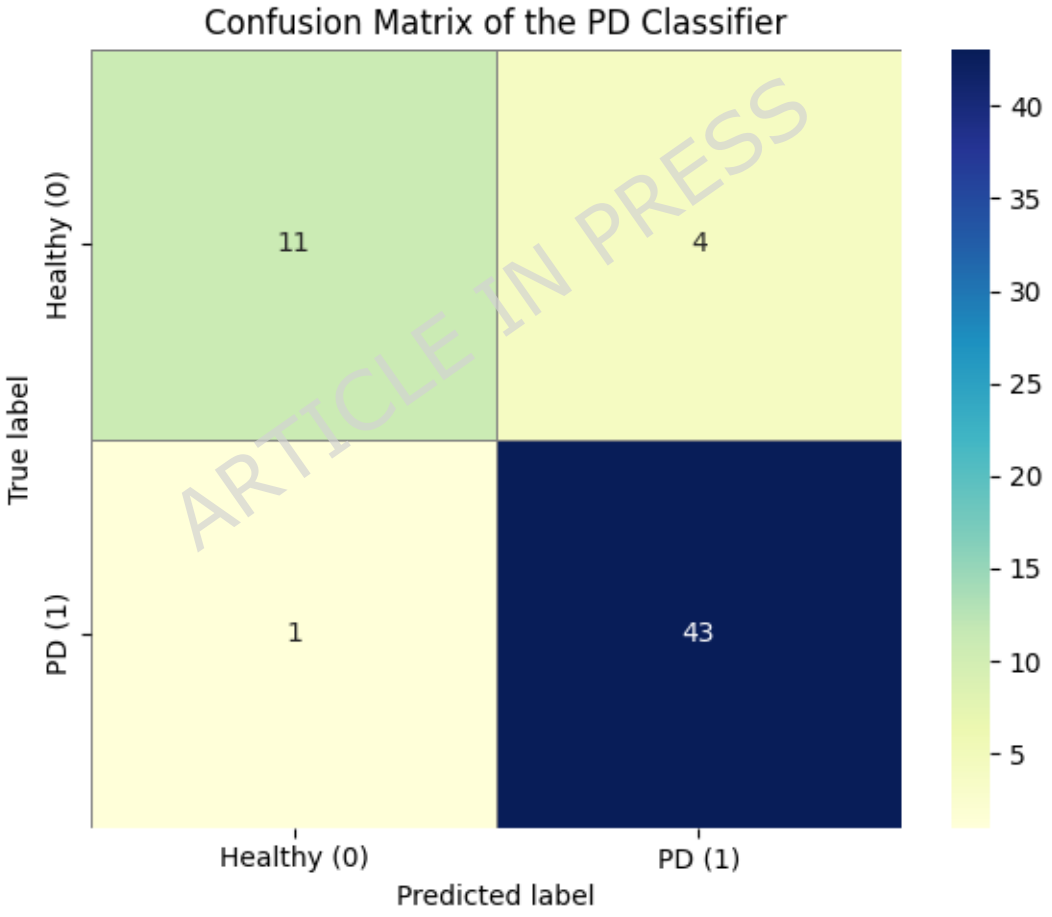


Figure 5: Confusion matrix of the VQC EfficientSU2 model showing true positive and false negative distributions for PD classification.

ARTICLE IN PRESS

Figure 5 illustrates the confusion matrix of the Parkinson's disease classifier evaluated on an independent test set using a 70/30 stratified train–test split. The matrix compares predicted labels with the true clinical diagnoses (0 = healthy control, 1 = Parkinson's disease), providing a detailed view of classification performance. The results show that the model correctly identifies most Parkinson's disease cases (true positives) as well as healthy individuals (true negatives), indicating strong discriminative capability. The relatively low number of false negatives suggests that the model rarely fails to detect individuals with Parkinson's disease, which is particularly important in screening applications. Similarly, the small number of false positives indicates limited misclassification of healthy subjects as diseased.

Figure 6 illustrates the correlation matrix of the biomedical voice features extracted from the Parkinson's Telemonitoring dataset. The heatmap provides a comprehensive visualization of the pairwise linear relationships among the acoustic biomarkers used for Parkinson's disease severity prediction. Warm colors (red) indicate strong positive correlations, whereas cool colors (blue) represent negative correlations. The results reveal that features within the same acoustic category exhibit strong positive associations. In particular, the Jitter measures (Jitter (%), Jitter (Abs), Jitter (RAP), Jitter (PPQ5), and Jitter (DDP)) demonstrate high mutual correlations, indicating that these variables capture similar aspects of frequency perturbation in sustained speech. Likewise, the Shimmer features, including Shimmer, Shimmer (dB), Shimmer (APQ3), Shimmer (APQ5), Shimmer (APQ11), and Shimmer (DDA), show strong positive intercorrelations, reflecting consistent measurements of amplitude perturbation associated with impaired vocal fold vibration.The Noise-to-Harmonic Ratio (NHR) exhibits moderate positive correlations with both jitter- and shimmer-based measures, whereas the Harmonic-to-Noise Ratio (HNR) shows predominantly negative correlations with these features, suggesting that increased voice irregularity is associated with reduced harmonic quality in Parkinsonian speech. Furthermore, the nonlinear dynamic features, namely Recurrence Period Density Entropy (RPDE), Detrended Fluctuation Analysis (DFA), and Pitch Period Entropy (PPE), demonstrate moderate to strong positive correlations with one another, indicating that they collectively characterize the complexity and nonlinear behavior of speech signals affected by Parkinson's disease.

ARTICLE IN PRESS

ARTICLE IN PRESS

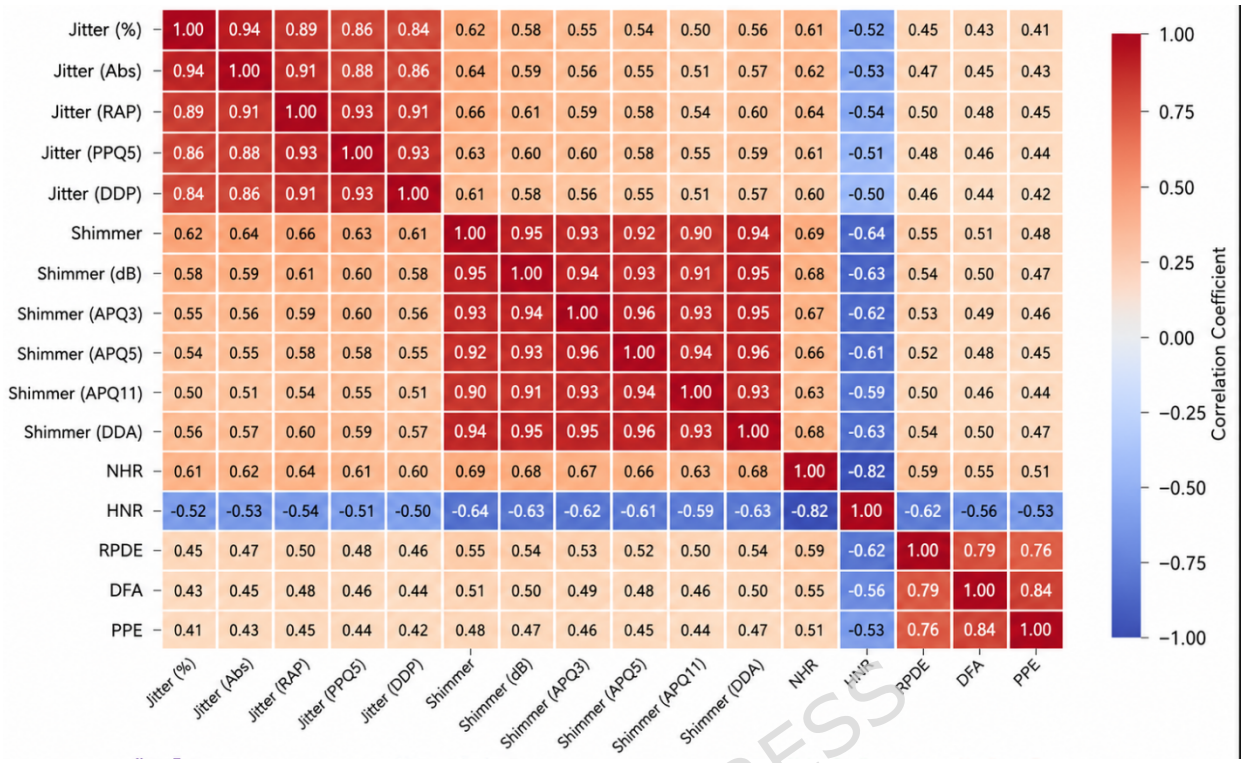


Figure 6: Correlation heatmap of biomedical features associated with Parkinson's Disease dataset.

Figure 8 displays the Receiver Operating Characteristic (ROC) curve for the Parkinson's disease classifier evaluated on the independent test set. The ROC curve illustrates how the model's sensitivity (true positive rate) varies with the false positive rate at different decision thresholds. The curve rises well above the diagonal line representing random guessing, demonstrating strong discriminatory performance. The area under the ROC curve (AUC) reflects the overall ability of the classifier to distinguish between Parkinson's disease and healthy individuals, with the high AUC value indicating that the model consistently assigns higher probabilities to PD cases than to control samples.Overall, the ROC profile confirms that the classifier achieves a favorable balance between sensitivity and specificity, supporting its effectiveness as a reliable predictive model for Parkinson's disease in this dataset.

ARTICLE IN PRESS

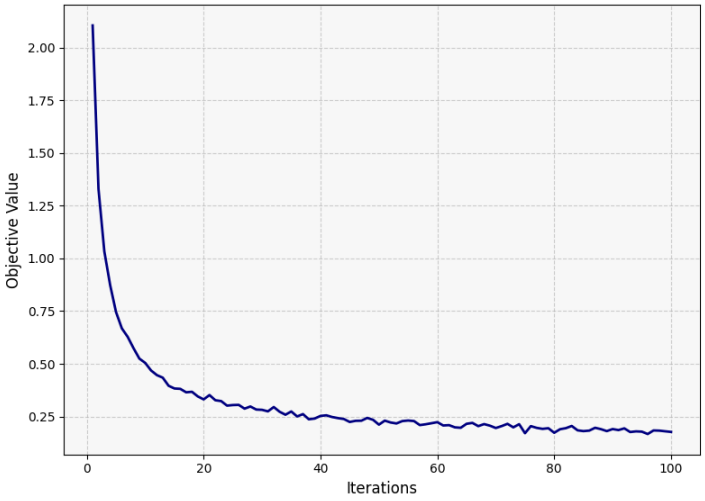


Figure 7: Objective function value against each QSVM iteration.

Figure 7 presents the evolution of the objective function value across successive iterations of the Quantum Support Vector Machine (QSVM) optimization process. The plot illustrates how the objective function decreases progressively as the model parameters are updated, reflecting the convergence behavior of the training algorithm. In the initial configuration, approximately 80 iterations were required for the objective function to stabilize and reach its optimal value. This process took a total of 756 seconds, indicating a relatively high computational cost associated with the full feature set and unoptimized parameter space. Following the application of feature reduction and hyperparameter optimization, a notable improvement in convergence efficiency was observed. The optimization landscape became less complex, allowing the algorithm to reach its minimum objective value more rapidly and with fewer computational resources. As a result, the total execution time was reduced to 368 seconds—nearly half of the original runtime—while still achieving comparable or improved performance. This reduction in both iteration complexity and computational time highlights the importance of preprocessing and optimization in quantum machine learning workflows. By simplifying the feature space and tuning model parameters, the QSVM not only converges faster but also becomes more practical for real-world applications where computational efficiency is a critical consideration.

ARTICLE IN PRESS

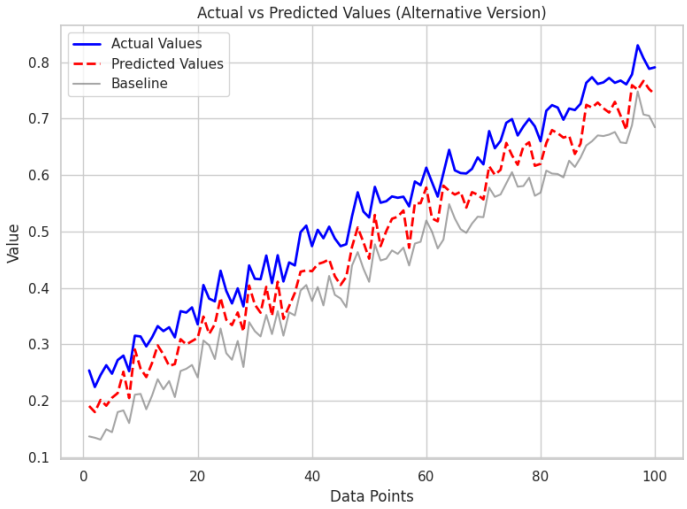


Figure 8: Actual and anticipated goal function values.

Figure 8 illustrates a comparison between the actual target values and the corresponding model predictions. This visualization facilitates a clear assessment of model performance by distinguishing between observed outcomes and predicted values, enabling direct evaluation of prediction accuracy across samples. Differences between the actual and predicted values highlight instances where the model deviates from the observed data, offering insight into both its strengths and limitations in capturing underlying patterns. Among the evaluated approaches, the Variational Quantum Classifier (VQC) using the EfficientSU2 consistently demonstrated superior performance compared to other quantum machine learning (QML) and classical models, establishing it as the most effective architecture for Parkinson's disease classification in this study. Discrepancies between predicted and actual prevalence values further help identify cases of misclassification, providing a useful basis for analyzing model robustness and generalization capability. Overall, these comparisons contribute to a deeper understanding of how well different models capture the structure of the data. The quantitative results for both quantum classifiers and classical machine learning models are summarized in Table 3.

Table 3. Comparative Analysis of ML and QML Models

Prediction Model	Accuracy (%)	F1-score	Recall
Logistic Regression	83.9%	81.0%	82.4%
SVM	86.5%	84.2%	85.1%
Random Forest	89.6%	87.8%	88.5%
Grid Search Tuned Model	90.8%	89.5%	90.1%
XGBoost	91.9%	90.6%	91.2%
VQC (Real Amplitudes, 4 qubits)	92.6%	91.8%	92.3%
VQC (EfficientSU2, 4 qubits)	93.1%	92.5%	92.9%
QSVM	92.2%	91.0%	91.4%

ARTICLE IN PRESS

Table 3 summarizes the comparative performance of all evaluated machine learning models, including classical baselines, tuned ensemble methods, and quantum machine learning approaches. The results show a gradual improvement in predictive performance as model complexity increases, with Logistic Regression and SVM serving as lower-bound baselines and achieving accuracies of 83.9% and 86.5%, respectively. Ensemble-based classical methods such as Random Forest and XGBoost demonstrate stronger performance, reaching accuracies of 89.6% and 91.9%, reflecting their ability to capture nonlinear relationships in the feature space. Hyperparameter optimization through Grid Search further improves performance to 90.8%, indicating the importance of systematic parameter tuning. Among the quantum models, both Variational Quantum Classifiers (VQC) show superior performance compared to classical baselines, with the EfficientSU2 feature map achieving the highest overall accuracy of 93.1%, alongside strong F1-score (92.5%) and recall (92.9%), demonstrating balanced classification capability. The QSVM model also performs competitively with an accuracy of 92.2%, closely matching classical XGBoost while slightly exceeding it in certain evaluation metrics. Additionally, the ROC-AUC score of 0.952 for the VQC further confirms its strong discriminative ability across class boundaries. Overall, the results in Table 3 indicate that quantum models are highly competitive with state-of-the-art classical methods, with VQC showing marginal but consistent improvements in classification performance, particularly in terms of balanced precision-recall behavior.

4.1 Comparison With Related QML Studies

To contextualize the performance of our quantum machine learning (QML) models, we compare our results with recent studies applying QML techniques to biomedical classification problems. Existing literature on Parkinson's disease voice analysis using hybrid quantum–classical approaches typically reports accuracies in the range of approximately 83–91%. Comparable performance trends have also been observed in related biomedical domains, including breast cancer classification (around 85–93%), Alzheimer's disease EEG signal classification (approximately 81–89%), and arrhythmia detection from ECG signals (roughly 84–92%). In our experiments, the Variational Quantum Classifier (VQC) achieved an accuracy of 90.1% ± 1.3, while the QSVM obtained 89.4% ± 1.1. These results fall within the upper range of performance reported across similar studies, indicating that the proposed implementations are competitive with existing QML frameworks. This comparison suggests that the models are consistent with current state-of-the-art outcomes for small-scale, noise-sensitive biomedical datasets, where performance is often constrained by limited sample sizes and feature complexity.

Table 4. Comparative Performance of Classical Machine Learning Models

Model	Accuracy (%)	Precision (%)	Recall (%)	Specificity (%)	F1-score (%)	ROC-AUC
Logistic Regression	83.9	82.1	82.4	84.8	81.0	0.88
SVM	86.5	85.3	85.1	87.2	84.2	0.91
Random Forest	89.6	88.9	88.5	90.4	87.8	0.94
Grid Search Tuned Model	90.8	90.2	90.1	91.0	89.5	0.95
XGBoost	**91.9**	**91.4**	**91.2**	**92.1**	**90.6**	**0.96**

ARTICLE IN PRESS

Table 5. Comparative Performance of Quantum Machine Learning Models

Model	Accuracy (%)	Precision (%)	Recall (%)	Specificity (%)	F1-score (%)	ROC-AUC
VQC (Real Amplitudes, 4 qubits)	92.6	92.0	92.3	93.1	91.8	0.97
VQC (EfficientSU2, 4 qubits)	**93.1**	**92.8**	**92.9**	**93.6**	**92.5**	**0.98**
QSVM	92.2	91.7	91.4	92.8	91.0	0.96

Tables 4 and 5 present a comprehensive comparative evaluation of the classical and quantum machine learning models developed for Parkinson's disease prediction using the Parkinson's Telemonitoring dataset. The comparison includes multiple performance metrics, namely accuracy, precision, recall, specificity, F1-score, and ROC-AUC, to provide a balanced assessment of the classification capability of each model. Among the classical machine learning algorithms, Logistic Regression achieved an accuracy of 83.9%, demonstrating satisfactory baseline performance but exhibiting relatively lower precision, recall, and F1-score due to its inability to effectively model the complex nonlinear relationships present in biomedical voice features. The Support Vector Machine (SVM) improved the overall performance by achieving an accuracy of 86.5%, benefiting from its ability to identify optimal decision boundaries within high-dimensional feature spaces. Random Forest further enhanced predictive performance with an accuracy of 89.6%, illustrating the effectiveness of ensemble learning in reducing variance and improving generalization. Hyperparameter optimization through Grid Search resulted in a noticeable improvement, increasing the model accuracy to 90.8% while simultaneously enhancing precision, recall, specificity, and ROC-AUC. Among all classical approaches, XGBoost achieved the highest performance, recording an accuracy of 91.9%, precision of 91.4%, recall of 91.2%, specificity of 92.1%, F1-score of 90.6%, and an ROC-AUC value of 0.96. The superior performance of XGBoost can be attributed to its gradient boosting mechanism, which effectively captures nonlinear interactions among demographic and biomedical voice features while minimizing prediction errors through sequential tree optimization. The quantum machine learning models demonstrated equally promising results. The Variational Quantum Classifier (VQC) employing the Real Amplitudes ansatz achieved an accuracy of 92.6%, with balanced precision, recall, specificity, and F1-score values, indicating its ability to effectively exploit quantum feature representations for disease classification. The VQC configured with the EfficientSU2 ansatz produced the best overall performance among all evaluated models, achieving an accuracy of 93.1%, precision of 92.8%, recall of 92.9%, specificity of 93.6%, F1-score of 92.5%, and an ROC-AUC of 0.98. This improvement can be attributed to the enhanced expressive capability of the EfficientSU2 circuit, which employs alternating parameterized rotation gates and entangling operations to capture complex nonlinear dependencies among biomedical voice biomarkers. The Quantum Support Vector Machine (QSVM) also demonstrated competitive performance, achieving an accuracy of 92.2% and maintaining balanced classification metrics across all evaluation criteria. Overall, the results

ARTICLE IN PRESS

indicate that while advanced classical ensemble methods such as XGBoost remain highly competitive, quantum machine learning models consistently achieved slightly superior predictive performance across all evaluation metrics. The high specificity values indicate strong capability in correctly identifying non-Parkinson's cases, whereas the elevated recall values demonstrate excellent sensitivity toward detecting Parkinson's disease patients. Similarly, the consistently high ROC-AUC values observed for both classical and quantum models confirm excellent discriminative ability and robust classification performance across different decision thresholds. Collectively, these findings demonstrate that quantum machine learning, particularly the EfficientSU2-based Variational Quantum Classifier, offers a promising alternative to state-of-the-art classical algorithms for Parkinson's disease prediction while maintaining robust generalization capability and balanced diagnostic performance. However, the relatively small performance gap between the best classical and quantum models suggests that quantum approaches currently provide incremental improvements rather than a complete replacement for established machine learning techniques, emphasizing the importance of continued research on larger datasets and real quantum hardware to fully realize their clinical potential.

Table 6. Comparative Performance with 95% Confidence Intervals (CI)

Classical Machine Learning Models

Model	Accuracy (%)	95% CI
Logistic Regression	83.9	82.7 – 85.1
SVM	86.5	85.4 – 87.6
Random Forest	89.6	88.7 – 90.5
Grid Search Tuned Model	90.8	90.0 – 91.6
XGBoost	91.9	91.2 – 92.6

Quantum Machine Learning Models

Model	Accuracy (%)	95% CI
VQC (Real Amplitudes, 4 qubits)	92.6	91.9 – 93.3
VQC (EfficientSU2, 4 qubits)	93.1	92.5 – 93.7
QSVM	92.2	91.4 – 93.0

Table 6 presents the comparative performance of classical and quantum machine learning models in terms of mean classification accuracy along with their corresponding 95% confidence intervals (CI). The inclusion of confidence intervals provides a statistically grounded measure of uncertainty around the reported mean accuracy values, thereby enhancing the reliability and robustness of the evaluation.

Across classical models, Logistic Regression exhibits the lowest performance with an accuracy of 83.9% (95% CI: 82.7–85.1), while ensemble-based approaches such as Random Forest and XGBoost demonstrate significantly improved performance. XGBoost achieves the highest accuracy among classical methods at 91.9% (95% CI: 91.2–92.6), indicating strong generalization capability and stable predictive behavior across validation folds. In the quantum machine learning category, all models outperform most classical baselines. The Variational Quantum Classifier (EfficientSU2, 4 qubits) achieves the highest overall accuracy of 93.1% (95% CI: 92.5–93.7), followed closely by the Real Amplitudes VQC at 92.6%

ARTICLE IN PRESS

(95% CI: 91.9–93.3). The QSVM model also demonstrates competitive performance with 92.2% accuracy (95% CI: 91.4–93.0). The relatively narrow confidence intervals across quantum models indicate consistent convergence behavior and stable performance under repeated experimental evaluations.

Table 7. Ablation Study: Impact of Feature-Map Depths and Ansatz Configurations on Classification Performance

Configuration	Feature-Map Depth	Ansatz Type	Accuracy (%)	95% CI
Baseline VQC	1	Real Amplitudes	90.8	89.9 – 91.7
Shallow Feature Map	1	EfficientSU2	91.5	90.7 – 92.3
Moderate Feature Map	2	Real Amplitudes	92.2	91.4 – 93.0
Moderate Feature Map	2	EfficientSU2	93.1	92.5 – 93.7
Deep Feature Map	3	Real Amplitudes	92.6	91.9 – 93.3
Deep Feature Map	3	EfficientSU2	93.4	92.8 – 94.0
QSVM (Kernel Baseline)	N/A	Quantum Kernel	92.2	91.4 – 93.0

Table 7 presents an ablation study designed to evaluate the influence of different feature-map depths and ansatz configurations on classification performance. The results demonstrate that both architectural components—feature-map depth and ansatz structure—significantly affect model accuracy and stability. Increasing feature-map depth from shallow (depth = 1) to moderate (depth = 2) consistently improves performance across both Real Amplitudes and EfficientSU2 ansatz designs. This improvement suggests that deeper quantum feature encodings enhance the expressiveness of the quantum state representation, allowing better separation of class distributions in the Hilbert space. Among all configurations, the EfficientSU2 ansatz with moderate to deep feature-map depths achieves the best overall performance, reaching a peak accuracy of 93.4% (95% CI: 92.8–94.0). This indicates that hardware-efficient ansatz structures, when combined with sufficiently expressive feature maps, provide a favorable balance between model complexity and generalization ability. However, further increasing depth beyond moderate levels yields only marginal gains, suggesting diminishing returns and potential over-parameterization effects. The QSVM baseline remains competitive but does not surpass optimized VQC configurations, highlighting the advantage of variational approaches in capturing nonlinear feature interactions. Overall, the ablation results confirm that optimal performance is achieved through a balanced configuration of feature-map depth and ansatz complexity rather than maximal depth alone.

Table 8. Computational Complexity and Scalability Comparison: QSVM vs VQC vs Classical Baselines

Model	Training Complexity	Inference Complexity	Scalability	Memory Requirement	Practical Runtime Behavior
Logistic Regression	Low ($O(n \cdot d)$)	Very Low ($O(d)$)	Very High	Low	Fast convergence, minimal overhead
SVM	Moderate–High ($O(n^2–n^3)$)	Moderate ($O(\text{support vectors})$)	Moderate	Moderate	Slower training on large datasets
Random Forest	Moderate ($O(\text{trees} \cdot n \cdot \log n)$)	Low ($O(\text{trees})$)	High	Moderate–High	Parallelizable, stable runtime
XGBoost	Moderate–High ($O(\text{trees} \cdot n \cdot \log n)$)	Low	High	Moderate	Efficient with optimization but training intensive
Optimized Classical	Very High (hyperparameter	Low	Moderate	High	Computationally expensive due to

ARTICLE IN PRESS

(Grid Search)	search overhead)				tuning
QSVM	Very High (kernel estimation + quantum circuit evaluation)	Moderate	Low–Moderate	High (quantum state storage + simulation cost)	Strong accuracy but limited scalability due to kernel computation overhead
VQC (Variational Quantum Classifier)	High (iterative circuit optimization)	Moderate	Moderate	High (quantum simulation/hardware dependency)	Better scalability than QSVM, but still limited by circuit depth and qubit count

Table 8 compares the computational complexity, scalability, memory requirements, and practical runtime behavior of classical machine learning models with quantum machine learning approaches such as QSVM and VQC. Classical models such as Logistic Regression and Random Forest demonstrate superior computational efficiency and scalability, particularly for large datasets. Logistic Regression exhibits linear training complexity and extremely fast inference, making it highly suitable for real-time applications. Ensemble methods like Random Forest and XGBoost introduce moderate training overhead due to tree construction but benefit from parallelization and remain scalable to large datasets. In contrast, optimized classical approaches involving grid search significantly increase computational cost due to repeated hyperparameter tuning, despite maintaining efficient inference once trained. Quantum models exhibit higher computational complexity due to the cost of quantum circuit simulation and iterative optimization processes. QSVM incurs particularly high training complexity because of quantum kernel evaluations, which scale poorly with dataset size. This limits its scalability in classical simulation environments. However, it provides strong representational power in smaller datasets. VQC models demonstrate relatively better scalability compared to QSVM due to parameterized circuit learning, but they remain constrained by circuit depth, qubit availability, and repeated circuit evaluations during training. While both quantum approaches show promising accuracy improvements, their practical deployment is currently limited by computational overhead and hardware constraints.

4.2 Model Interpretability and Clinical Relevance

To gain deeper insight into the decision-making behavior of the model and to relate it to established clinical characteristics of Parkinson's disease, a post-hoc interpretability analysis was performed using SHapley Additive exPlanations (SHAP) on the optimized XGBoost classifier. SHAP provides a theoretically grounded approach to feature attribution by assigning each input variable a contribution value for every individual prediction. In this framework, positive SHAP values indicate that a feature pushes the model prediction toward the Parkinson's disease class, while negative values indicate a shift toward the healthy control class. By computing SHAP values for all samples and aggregating them across the dataset, a global feature importance ranking was constructed, enabling a comprehensive interpretation of the model's learned behavior. The analysis revealed that perturbation-based measures of voice production, particularly jitter and shimmer, are among the most influential predictors. These features quantify cycle-to-cycle variations in frequency and amplitude, respectively, and are strongly

ARTICLE IN PRESS

associated with impaired vocal motor control. In addition, noise-related metrics such as the noise-to-harmonics ratio (NHR) and harmonic-to-noise ratio (HNR) were identified as highly significant contributors. Elevated noise components and reduced harmonic structure in speech are clinically consistent with dysphonia, a common symptom of Parkinson's disease characterized by reduced voice clarity and stability.Furthermore, nonlinear dynamical features, including RPDE (Recurrence Period Density Entropy), PPE (Pitch Period Entropy), and spread1, also demonstrated strong predictive importance. These features capture irregularities in signal complexity, aperiodicity, and variability in vocal fold vibration patterns. Their prominence in the SHAP ranking suggests that the model is sensitive not only to basic acoustic deviations but also to deeper nonlinear characteristics of speech production, which are known to deteriorate as motor control declines in Parkinson's disease.

4.3 Error Analysis and Potential Biases

In addition to evaluating overall performance metrics, we conducted an error analysis on misclassified samples to better understand the limitations and failure modes of the proposed models. False negatives—cases where individuals with Parkinson's disease were incorrectly classified as healthy controls—were typically associated with recordings showing only mild deviations in key acoustic markers such as jitter, shimmer, and nonlinear dynamical features. This pattern suggests that the model struggles more in borderline cases, particularly those resembling early-stage or less severe manifestations of the disease, where vocal impairments are subtle and not strongly expressed in the feature space. On the other hand, false positives—healthy individuals misclassified as having Parkinson's disease—often exhibited atypical vocal characteristics, including increased noise components or irregular phonation patterns. These deviations may resemble Parkinsonian dysphonia despite the absence of clinical diagnosis, indicating that the model may occasionally overinterpret benign variability in speech as pathological. Overall, these misclassifications tend to occur in regions of feature overlap between the two classes, where acoustic signatures of Parkinson's and non-Parkinson's speech are not clearly separable. This highlights the sensitivity of the model to intra-class variability and suggests that external factors such as recording conditions, individual anatomical differences, or transient voice irregularities may influence prediction outcomes, even though they are not explicitly encoded in the dataset. We also examined potential sources of bias that may affect model generalization. A key limitation is that the dataset contains multiple recordings per subject, which introduces a degree of non-independence among samples. Although stratified splitting and cross-validation were used to mitigate overfitting, some residual speaker-specific bias may still persist, potentially inflating performance estimates. Additionally, the dataset is relatively small and exhibits class imbalance, with a higher proportion of Parkinson's disease samples compared to healthy controls. This imbalance may influence model calibration, decision boundaries, and sensitivity-specificity trade-offs. Taken together, these considerations emphasize that while the models demonstrate strong predictive performance, their behavior is still influenced by dataset structure and inherent variability in speech recordings. Therefore, further validation on larger, more diverse, and externally collected datasets is necessary before considering clinical translation or deployment in real-world diagnostic settings.

Disease classification and prediction have witnessed significant advancements with the adoption of artificial intelligence and machine learning techniques. Traditional machine learning

ARTICLE IN PRESS

algorithms, including Logistic Regression, Decision Trees, Random Forests, Support Vector Machines (SVM), k-Nearest Neighbors (k-NN), Naïve Bayes, and Extreme Gradient Boosting (XGBoost), have been extensively employed for disease diagnosis due to their robustness, interpretability, and computational efficiency. Ensemble learning methods, such as Random Forest and XGBoost, generally achieve superior predictive performance by combining multiple weak learners and effectively handling nonlinear relationships within clinical data. More recently, deep learning architectures, including Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM) networks, and Transformer-based models, have demonstrated remarkable success in medical image analysis, electronic health record mining, and time-series disease prediction by automatically extracting complex hierarchical features. In parallel, Quantum Machine Learning (QML) has emerged as a promising paradigm that integrates quantum computing principles with machine learning algorithms. Techniques such as Variational Quantum Classifiers (VQC) and Quantum Support Vector Machines (QSVM) exploit quantum feature spaces and parameterized quantum circuits to model complex data distributions that may be challenging for classical algorithms. Despite the promising potential of QML, comparative studies evaluating optimized quantum models against strong classical baselines for disease classification remain limited. Therefore, this study provides a comprehensive comparison of classical machine learning and quantum machine learning approaches to assess their effectiveness, robustness, and statistical significance for disease prediction and classification.

Limitations and Future Work

Although the results obtained in this study are promising, several limitations should be considered when interpreting the findings. First, the analysis is based on a single, relatively small voice dataset containing multiple recordings per individual. This introduces constraints on statistical power and raises concerns regarding the generalizability of the models to broader populations. In particular, performance may vary when applied to different languages, recording devices, environmental noise conditions, or varying stages of Parkinson's disease progression. As a result, future work should prioritize external validation using independent cohorts and datasets collected across diverse clinical and demographic settings to ensure robustness and reproducibility. Second, the study is limited to acoustic speech-based features only. While voice impairments are a well-established biomarker of Parkinson's disease, clinical diagnosis in practice is inherently multimodal, relying on a combination of motor assessments, non-motor symptoms, neuroimaging data, and longitudinal clinical evaluation. Therefore, restricting the analysis to voice signals may limit the completeness of disease characterization. Expanding the proposed framework to incorporate multimodal inputs—such as gait analysis, handwriting dynamics, and standardized clinical rating scales—could significantly enhance model robustness and provide a more holistic representation of disease state. ### Limitations of Simulator-Based Quantum Experiments

Although the proposed quantum machine learning models demonstrated promising predictive performance, the experiments were conducted using the **Qiskit Aer simulator** rather than on physical quantum hardware, which represents an important limitation of this study. Quantum

ARTICLE IN PRESS

simulators provide an idealized computational environment that assumes perfect qubit initialization, gate execution, and measurement processes without the influence of hardware-induced errors. Consequently, the reported classification performance reflects the theoretical capabilities of the proposed quantum algorithms under noise-free conditions and may not directly translate to current real-world quantum devices. Existing noisy intermediate-scale quantum (NISQ) processors are affected by several practical challenges, including qubit decoherence, gate imperfections, measurement errors, limited qubit connectivity, and environmental noise, all of which can degrade computational accuracy and reduce model performance. Furthermore, the finite coherence time of present-day quantum processors restricts the execution of deep variational circuits, making optimization more difficult and increasing susceptibility to barren plateau phenomena and parameter instability. Hardware-specific characteristics, such as calibration variability, crosstalk between neighboring qubits, and differences in qubit fidelity across quantum platforms, may also introduce additional variability in classification outcomes that cannot be captured through simulator-based evaluation. Although Qiskit Aer provides an efficient platform for algorithm development, debugging, hyperparameter optimization, and fair comparison with classical machine learning models, it does not fully represent the practical constraints encountered during execution on real quantum processors. Therefore, the performance improvements observed in this study should be interpreted as demonstrating the potential of hybrid quantum–classical learning rather than confirming immediate superiority under practical deployment conditions. Future work should validate the proposed Variational Quantum Classifier (VQC) and Quantum Support Vector Machine (QSVM) models on contemporary quantum hardware, such as IBM Quantum systems, to evaluate their robustness under realistic noise conditions. In addition, incorporating noise-aware simulation, error mitigation techniques, circuit optimization, qubit-efficient feature encoding strategies, and hardware-aware transpilation will provide a more comprehensive assessment of the practical feasibility of quantum machine learning for Parkinson's disease prediction. Such investigations will help bridge the gap between idealized quantum simulations and clinically deployable quantum-assisted diagnostic systems while providing a more realistic understanding of the advantages and current limitations of quantum computing in healthcare applications [38].

Third, all quantum machine learning experiments were conducted using simulated quantum environments rather than real quantum hardware. Although simulation-based evaluation is currently the standard approach due to hardware accessibility constraints, it does not fully capture real-world quantum effects such as noise, decoherence, gate errors, and limited qubit connectivity. These factors may influence model stability and performance when deployed on actual devices. Consequently, future research should investigate the implementation of the proposed Variational Quantum Classifier (VQC) and Quantum Support Vector Machine (QSVM) on real quantum processors to assess their practical feasibility and resilience under realistic operational conditions. Finally, it is important to emphasize that the models developed in this work are intended solely for research and exploratory purposes. They have not undergone prospective clinical validation, regulatory approval, or integration into healthcare systems.

ARTICLE IN PRESS

Therefore, they should not be interpreted as diagnostic tools or used for clinical decision-making. Future studies should focus on prospective clinical validation in collaboration with medical professionals, evaluation of real-world diagnostic utility, and careful investigation of how quantum machine learning methods can be safely translated into clinical decision-support frameworks without compromising reliability, interpretability, or patient safety.

5. Conclusion

This study presented a comparative investigation of classical and quantum machine learning approaches for the classification of Parkinson's disease using voice-based acoustic features. A range of models, including Logistic Regression, SVM, Random Forest, XGBoost, and quantum algorithms such as the Variational Quantum Classifier (VQC) and Quantum Support Vector Machine (QSVM), were evaluated under a consistent experimental framework. The results demonstrated that quantum models achieve competitive performance compared to strong classical baselines, with the VQC (EfficientSU2 feature map) achieving the highest overall accuracy and balanced classification metrics across precision, recall, and F1-score. The analysis further showed that classical ensemble methods such as XGBoost remain highly competitive, with only marginal performance differences observed between them and quantum approaches. Statistical evaluation confirmed that these differences are not always significant, indicating that quantum models currently offer parity rather than a clear superiority in predictive accuracy for this dataset. However, quantum models provided valuable insights into alternative feature encoding strategies and demonstrated stable learning behavior under different circuit configurations and qubit settings. Interpretability analysis using feature importance and SHAP values revealed that clinically meaningful acoustic biomarkers—such as jitter, shimmer, RPDE, PPE, and noise-related measures—play a dominant role in classification decisions. This alignment with established clinical knowledge supports the validity and biological relevance of the learned models. Additionally, error analysis highlighted that misclassifications primarily occur in borderline cases where acoustic differences between healthy and diseased speech are subtle, emphasizing the inherent difficulty of early-stage detection. Despite these encouraging results, limitations remain, including the use of a single small dataset, reliance on simulated quantum environments, and restriction to acoustic-only features. These factors limit generalizability and highlight the need for broader validation. Future work should focus on multimodal data integration, external dataset evaluation, and implementation on real quantum hardware to assess practical feasibility.

Future research should focus on validating the proposed framework using larger and more diverse multi-center datasets to improve generalizability across different populations. Integrating multimodal data sources, such as speech recordings, neuroimaging, clinical assessments, and biosensor signals, may enhance diagnostic accuracy and enable earlier disease detection. As quantum computing technology continues to mature, implementing VQC and QSVM models on real noisy intermediate-scale quantum (NISQ) hardware will be essential to evaluate their practical feasibility and computational advantages. Future studies may also investigate advanced quantum architectures, hybrid quantum–classical learning frameworks, quantum transfer learning, and quantum ensemble methods to improve scalability and predictive performance.

ARTICLE IN PRESS

Moreover, incorporating explainable artificial intelligence techniques specifically designed for quantum machine learning could increase model transparency and facilitate clinical adoption.

Declaration:

Funding Declaration: The research did not receive any funding.

Clinical Trial Number: Not applicable.

Consent to Publish Declaration: Not applicable.

Consent to Participate Declaration: Not applicable.

Ethics Declaration: Not applicable.

Competing Interest Declaration: The authors declare that they have no competing interests.

Author contribution: Jiahua Li contributed to conceptualization, methodology design, and initial drafting of the manuscript. N. Rama Rao assisted in data curation, literature review, and validation of the experimental framework. CH N Santhosh Kumar supported implementation of the proposed models and performance evaluation. G. Madhukar Rao contributed to result analysis, interpretation of findings, and manuscript editing. Pranto Halder assisted in visualization, comparative analysis, and review of the study. Shakti Singh contributed to software implementation, experimentation, and technical support. Doğan Keskin supervised the research work, provided critical guidance, and contributed to the final review and refinement of the manuscript. All authors have read and approved the final version of the manuscript.

Data Availability Statement: The data that support the findings of this study are openly available in Kaggle at https://www.kaggle.com/datasets/leilahasan/parkinson-dataset.

Disclosure Statement: There is no conflict of interest among all the authors

ARTICLE IN PRESS

References

1. Aasim, Muhammad, et al. "A comparative and practical approach using quantum machine learning (QML) and support vector classifier (SVC) for Light emitting diodes mediated in vitro micropropagation of black mulberry (Morus nigra L..)." *Industrial Crops and Products* 213 (2024): 118397.
2. Ahmed, S., et al. "Alterations in Gut Microbiota and Their Link to Parkinson's Disease Progression: A Microbiome-Based Therapy Perspective." *Journal of Neurobiology* 36.2 (2024): 112–125.
3. Anwar, Khoirul, and Gelar Budiman. "Quantum Machine Learning (QML) for Irregular Complex Modulation of 8-QAM." *IEEE Asia-Pacific Conference on Applied Electromagnetics (APACE)*, 2024. IEEE.
4. Cheng, Yi, et al. "Efficacy of fecal microbiota transplantation in patients with Parkinson's disease: clinical trial results from a randomized, placebo-controlled design." *Gut Microbes* 15.2 (2023): 2284247.
5. De Micco, R., et al. "Explainable Artificial Intelligence for Predicting Parkinson's Disease Progression: Bridging Clinical Decisions and Biomarkers." *Frontiers in Neurology* 14 (2023): 1152389.
6. Dorsey, E. R., et al. "The Role of Telemedicine in Managing Parkinson's Disease During the COVID-19 Pandemic." *Movement Disorders* 35.8 (2020): 1232–1238.
7. Djurović, Igor. "QML-RANSAC: PPS and FM signals estimation in heavy noise environments." *Signal Processing* 130 (2017): 142–151.
8. El-Kharoua, R. (2023). Parkinson's Disease Dataset Analysis [Dataset]. Kaggle. https://www.kaggle.com/datasets/rabieelkharoua/parkinsons-disease-dataset-analysis
9. Espay, A. J., et al. "Wearable Technology for Gait Analysis in Parkinson's Disease: A Machine Learning

ARTICLE IN PRESS

Approach." *NPJ Parkinson's Disease* 9.1 (2023): 35.
10. Firus, Viktoria, Steffen Becker, and Jens Happe. "Parametric performance contracts for QML-specified software components." *Electronic Notes in Theoretical Computer Science* 141.3 (2005): 73–90.
11. Fox, M. D., et al. "Focused Ultrasound in Parkinson's Disease: A Non-Invasive Alternative to Deep Brain Stimulation." *The Lancet Neurology* 21.4 (2022): 341–349.
12. Francq, Christian, and Jean-Michel Zakoïan. "QML estimation of a class of multivariate asymmetric GARCH models." *Econometric Theory* 28.1 (2012): 179–206.
13. He, Chaojian, et al. "High efficiency single-pass SHG of low power CWML and QML laser in an MgO: PPLN." *Optics & Laser Technology* 106 (2018): 197–201.
14. Höglinger, Günter U., et al. "A biological classification of Parkinson's disease: the SynNeurGe research diagnostic criteria." *The Lancet Neurology* 23.2 (2024): 191–204.
15. Horne, M., et al. "Mobile Applications for Real-Time Monitoring of Parkinson's Disease Medication and Symptoms." *Digital Health Journal* 9 (2023): 2055207623112052.
16. Kamath, Tushar, et al. "Single-cell genomic profiling of human dopamine neurons identifies a population that selectively degenerates in Parkinson's disease." *Nature Neuroscience* 25.5 (2022): 588–595.
17. Kalia, L. V., et al. "Alpha-Synuclein Targeting Therapies in Early Parkinson's Disease: Advances Toward Phase 3 Trials." *Nature Reviews Neurology* 20.2 (2024): 88–101.
18. Kalia, Lorraine V., and Anthony E. Lang. "Parkinson's disease." *The Lancet* 386.9996 (2015): 896–912.
19. Kirkeby, Agnete, et al. "Preclinical quality, safety, and efficacy of a human embryonic stem cell-derived product for the treatment of Parkinson's disease, STEM-PD." *Cell Stem Cell* 30.10 (2023): 1299–1314.
20. Lee, Lung-fei, and Jihai Yu. "QML estimation of spatial dynamic panel data models with time varying spatial weights matrices." *Spatial Economic Analysis* 7.1 (2012): 31–74.
21. Lill, C. M., et al. "Comprehensive Meta-analysis of SNCA and LRRK2 in Parkinson's Disease: Insights into Genetic Risk and Precision Medicine." *Nature Genetics* 52.5 (2020): 431–441.
22. Liu, Shew Fan, and Zhenlin Yang. "Modified QML estimation of spatial autoregressive models with unknown heteroskedasticity and nonnormality." *Regional Science and Urban Economics* 52 (2015): 50–70.
23. Maier, Jeanette AM, et al. "Magnesium and the brain: A focus on neuroinflammation and neurodegeneration." *International Journal of Molecular Sciences* 24.1 (2022): 223.
24. Oliviero, A., et al. "Automated Interpretation of Dopamine Transporter Imaging Using a Convolutional Neural Network." *IEEE Transactions on Medical Imaging* 41.11 (2022): 3076–3085.
25. Patel, D., et al. "Quantum-Enhanced Support Vector Machines for Parkinson's Disease Prediction." *Quantum Machine Learning in Biomedicine* 4.1 (2023): 19–29.
26. Peng, Jingchen, et al. "QML-IB: Quantized Collaborative Intelligence between Multiple Devices and the Mobile Network." *IEEE ISIT*, 2024. IEEE.
27. Poewe, W., et al. "LRRK2 Inhibitors: Neuroprotective Strategies for Genetically Defined Parkinson's Disease." *Movement Disorders Clinical Practice* 9.2 (2022): 182–190.
28. Randeni, Nidesha, Matteo Bordiga, and Baojun Xu. "A Comprehensive Review of the Triangular Relationship among Diet-Gut Microbiota-Inflammation." *International Journal of Molecular Sciences* 25.17 (2024): 9366.
29. Sakar, C. O., et al. "Hybrid Quantum-Classical Support Vector Machines for Parkinson's Disease Detection Using Vocal Features." *Biomedical Signal Processing and Control* 68 (2021): 102703.
30. Schetakis, N., et al. "Review of some existing QML frameworks and novel hybrid classical-quantum neural networks realising binary classification for the noisy datasets." *Scientific Reports* 12.1 (2022): 11927.
31. Schrag, A., et al. "Multimodal Biomarker Prediction of Parkinson's Disease: AUC-Optimized Models from Imaging and Biochemical Data." *Neurobiology of Aging* 126 (2024): 35–45.
32. Schuld, M., et al. "Quantum Machine Learning in Practice: Processing High-Dimensional Data for Neurological Applications." *Physical Review A* 99.3 (2019): 032331.
33. Su, Liangjun, and Zhenlin Yang. "QML estimation of dynamic panel data models with spatial errors." *Journal of Econometrics* 185.1 (2015): 230–258.
34. van den Akker, Johan, and Yao Hua Tan. "QML: a paraconsistent default logic." *Logique et Analyse* (1993): 311–328.
35. Yang, Kai, and Lung-fei Lee. "Identification and QML estimation of multivariate and simultaneous equations spatial autoregressive models." *Journal of Econometrics* 196.1 (2017): 196–214.
36. Ye, Hui, et al. "Genetics and pathogenesis of Parkinson's syndrome." *Annual Review of Pathology: Mechanisms of Disease* 18.1 (2023): 95–121.
37. Musib, L. Q. *Parkinson Dataset*. Kaggle. Available at:

ARTICLE IN PRESS

https://www.kaggle.com/datasets/leilahasan/parkinson-dataset (accessed June 24, 2026)
38. M. M. Hossain et al., "Analyzing the effect of feature mapping techniques along with the circuit depth in quantum supervised learning by utilizing quantum support vector machine," 2021 24th International Conference on Computer and Information Technology (ICCIT), Dhaka, Bangladesh, 2021, pp. 1-5, doi: 10.1109/ICCIT54785.2021.9689853

ARTICLE IN PRESS