NeuroSignal Precision: A Hierarchical Approach for Enhanced Insights in Parkinson's Disease Classification

1st Kazi Shaharair Sharif
Department of Computer Science
University of South Dakota
Vermillion, SD, USA
kazi.sharif@okstate.edu

2nd Mohammed Majbah Uddin
School of Business and Technology
Emporia State University
Emporia, Kansas, USA
muddin@g.emporia.edu

3rd Md Abubakkar
Department of Computer Science
Midwestern State University
Dallas, TX, USA
mabubakkar1116@my.msutexas.edu

***Abstract*—Parkinson's disease (PD) is a progressive neurological disorder that affects movement, posture, handwriting, and speech. Parkinson's disease is challenging to diagnose early due to subtle symptoms that often go unnoticed, necessitating reliable and accurate classification models to aid clinical decision-making. This research introduces a comprehensive benchmarking of nine unified models, and a unique contribution of this research is the adaptation of the Tabular Transformer model for structured medical data, achieving an unprecedented accuracy of 99.49%, setting a new benchmark for Parkinson's disease classification. The proposed approach provides an advanced, adaptable framework that supports clinicians in making early, accurate diagnoses, ultimately improving patient care. In contrast to previous studies that predominantly emphasize traditional models, this research employs attention-based deep learning to capture complex feature interactions, achieving substantially higher accuracy. The study evaluates nine models: SVM, Decision Tree, Random Forest, AdaBoost, Gradient Boosting, XGBoost, KNN, CNN, and Tabular Transformer, achieving improved accuracy across all models compared to previous studies, marking a notable advancement in Parkinson's disease classification performance. The Transformer's attention mechanism captures intricate data patterns, providing clear advantages over traditional approaches and improving diagnostic precision for early-stage Parkinson's detection. Data preprocessing included the Synthetic Minority Over-sampling Technique for class balancing and feature standardization, with each model, from SVM and Decision Trees to CNN and XGBoost, optimized through Optuna for optimal performance. This research offers the medical field a versatile, high-accuracy framework that aids clinicians in timely and reliable PD diagnosis, potentially improving patient outcomes and advancing Parkinson's diagnostic tools for future clinical use.**

***Index Terms*—Structured Medical Data, Transformer, Parkinson's disease, Hyperparameter optimization, Attention-layers**

I. Introduction

Parkinson's disease (PD) is a progressive neurodegenerative disorder that affects millions of people worldwide and is characterized by a complex array of motor and non-motor symptoms. The hallmark motor symptoms include tremors, rigidity, bradykinesia, and postural instability, while non-motor manifestations encompass cognitive decline, depression, and anxiety [1]. As the global population ages, the prevalence of PD is expected to rise significantly, presenting an increasing challenge to healthcare systems and society [2]. Early and accurate diagnosis of PD is crucial for effective management and improved patient outcomes. The initial symptoms of PD are often subtle and can overlap with those of other neurological conditions, complicating differential diagnosis and potentially leading to misdiagnosis or delayed treatment [3], [4]. In recent years, the application of advanced computational techniques, particularly machine learning (ML) and deep learning (DL), has shown promising results in supporting early PD detection. These approaches offer the ability to detect complex patterns within medical data that may be missed by traditional diagnostic methods [5]. The rapid advancements in ML and DL architectures, such as Convolutional Neural Networks (CNNs) and Transformers, have further enhanced the potential for capturing intricate data patterns, thus improving classification accuracy and robustness in PD diagnosis [6]. Our research encompasses a wide range of algorithmic approaches, from traditional classifiers to advanced deep-learning architectures. A key innovation in this study is the introduction of the Tabular Transformer, an attention-based deep learning model rarely applied in PD classification, which effectively captures complex feature interactions within structured data. By implementing and refining multiple models coupled with extensive hyperparameter optimization, this research seeks to establish a high-accuracy, adaptable framework that significantly enhances PD classification accuracy and robustness. The comprehensive nature of this study, combining diverse modeling approaches with novel methodological refinements, has the potential to make a substantial impact on PD diagnostics [7]. Ultimately, this research aims to provide healthcare professionals with an advanced diagnostic tool to support timely and reliable PD diagnoses. The advanced diagnostic tool developed could support timely, reliable PD diagnoses, facilitate early interventions, and improve care quality and outcomes, ultimately benefiting patients facing this challenging neurological disorder [8].

II. Literature Review

Recent advancements in machine learning (ML) and deep learning (DL) have significantly contributed to Parkinson's disease (PD) diagnostics, addressing the limitations of traditional methods by identifying complex patterns in medical data. PD is a multifaceted disorder with motor and non-motor symptoms that often overlap with other neurodegenerative conditions, complicating diagnosis, especially in the early stages [1], [9]. The need for reliable and accurate diagnostic tools has driven research toward computational approaches that improve early detection and patient outcomes [2], [4].

979-8-3315-0649-0/24/$31.00 ©2024 IEEE

ML techniques like SVM and Decision Trees have been extensively applied to PD classification due to their interpretability and relatively strong baseline performance. Srinivasan et al. (2024) [10] demonstrated the efficacy of SVM combined with Recursive Feature Elimination in PD detection from voice signals, achieving a high accuracy of 93.84%. However, the focus on voice data limits its application, whereas our study extends to a broader set of biomarkers. Singh et al. (2023) [11] also highlighted that ensemble methods like Random Forests and Gradient Boosting enhance accuracy over single classifiers. However, they may lack the flexibility needed for complex, early-stage PD diagnosis. DL approaches, particularly Convolutional Neural Networks (CNNs), have shown great promise in medical diagnostics, as they can capture spatial and sequential relationships in structured data. Camacho et al. (2023) [12] developed an explainable DL model using T1-weighted MRI data, achieving high accuracy in distinguishing PD patients from healthy controls. While our study also incorporates MRI data, we further expand by integrating multiple data modalities and employing a broader set of models, including non-image data, to enhance diagnostic insights. Thakur et al. (2022) [13] introduced a CNN model with a soft-attention mechanism for better feature extraction in DaTscan images, which aligns with our approach of utilizing advanced DL architectures. However, our work uniquely includes the Tabular Transformer, an attention-based model that effectively handles structured data for PD, capturing intricate feature interactions beyond the capabilities of CNNs alone.

Multimodal techniques have shown promise in PD diagnostics by incorporating diverse data types, such as imaging and clinical markers. Musti et al. (2024) [7] investigated multimodal DL architectures for prodromal PD detection, achieving a 95.5% accuracy by integrating various 3D models. While similar in concept, our study distinguishes itself by including both ML and DL models with an innovative application of the Tabular Transformer. Originally designed for language processing, the Transformer's ability to capture long-range dependencies makes it well-suited for structured data and differentiates our research from previous studies [14]. The Tabular Transformer has rarely been applied in PD research, making this one of the first studies to explore its potential for complex clinical data interpretation. Effective hyperparameter optimization is essential for maximizing model performance, particularly in complex medical diagnostics. While previous studies have incorporated automated optimization techniques, few have applied extensive, model-specific tuning across a wide array of models for PD classification. Our study utilizes Optuna for systematic hyperparameter optimization, a method demonstrated by Yoshida and Kameda (2021) and Kaur et al. (2022) [6], [15] to significantly improve accuracy and generalizability in ML applications. This research stands out by benchmarking nine diverse models for PD classification, including traditional ML, ensemble methods, and advanced deep learning architectures like CNNs and the Tabular Transformer. Unlike prior studies focused on limited model comparisons, our comprehensive evaluation establishes a robust framework for PD diagnostics. Additionally, our study addresses common diagnostic challenges, such as class imbalance and inadequate validation, through rigorous cross-validation and explicit class balancing. [16].

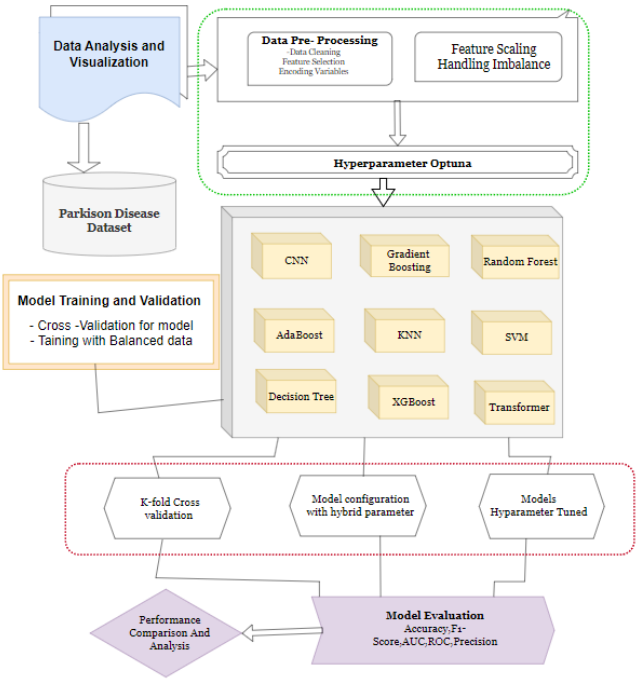

Fig. 1. WorkFlow Diagram

III. Methodology

Fig. 1 workflow diagram illustrates the end-to-end process of Parkinson's disease classification and was designed to explore and optimize various machine learning (ML) and deep learning (DL) models) using a structured data approach. Each step was meticulously crafted to address common challenges in PD data analysis, such as class imbalance and feature variability while leveraging advanced models and tuning processes to enhance diagnostic accuracy.

A. Overview of Data Collection

The dataset used in this study was obtained from the UCI Parkinson's Disease dataset, a well-structured dataset containing key vocal biomarkers for distinguishing PD patients from healthy controls. The dataset included features related to vocal fold function and speech, which have been shown to carry significant diagnostic potential in PD. Fig. 2 illustrates the imbalance in the dataset. To ensure unbiased results, we partitioned the dataset into training and test sets with an 80:20 split, maintaining a stratified approach to ensure a balanced representation of PD and control groups in both sets [17]. Fig. 3 presents the distributions of the most significant acoustic features used for classification, such as jitter, shimmer, fundamental frequency, and harmonic-to-noise ratio.

B. Data Preprocessing

The data preprocessing pipeline was designed to enhance model performance by addressing class imbalance and ensuring feature standardization, essential for robust Parkinson's disease (PD) classification [18]. Fig. 4 shows the correlations between features, identifying dependencies and redundancies within the dataset's acoustic measurement. Given the imbalance in the dataset, the Synthetic Minority Over-sampling Technique (SMOTE) was employed to generate synthetic samples for the minority class, effectively balancing the

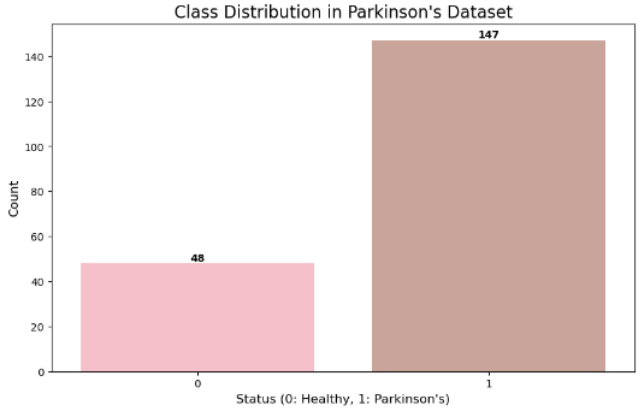

Fig. 2. Distribution Status

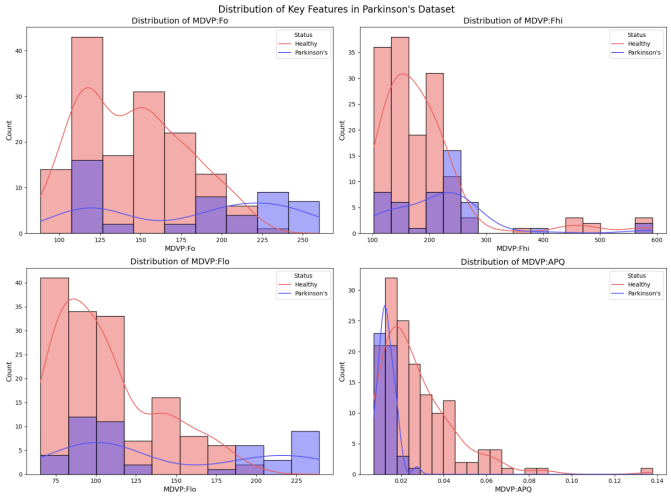

Fig. 3. Class and key features Distribution

distribution and reducing bias that could hinder model performance [19]. This approach is particularly crucial for medical datasets, where imbalances can skew predictions, and here, it allowed the models to capture patterns effectively across both PD-positive and control groups. To standardize feature scales, each feature was transformed using StandardScaler, adjusting all features to zero mean and unit variance. This step was indispensable, particularly for gradient-based models and deep learning architectures, which are sensitive to feature scaling and can benefit from consistent magnitudes during training [20]. A distinctive aspect of this preprocessing pipeline was its integration with Optuna-based hyperparameter tuning. Unlike conventional methods, which apply preprocessing independently, our approach incorporated SMOTE and scaling directly within the tuning process. This integration allowed models to leverage balanced, standardized data across each hyperparameter trial, optimizing performance through consistent data preparation.

C. Hyperparameter Optimization

Optuna, an advanced hyperparameter optimization framework, systematically tune parameters across all nine classifiers, enhancing performance and robustness in Parkinson's disease classification. Leveraging Optuna's efficient techniques, such as grid search and Bayesian optimization, enabled us to explore a broad spectrum of hyperparameters, including learning rate, batch size, neural network depth, and the number of estimators for ensemble models. This tuning

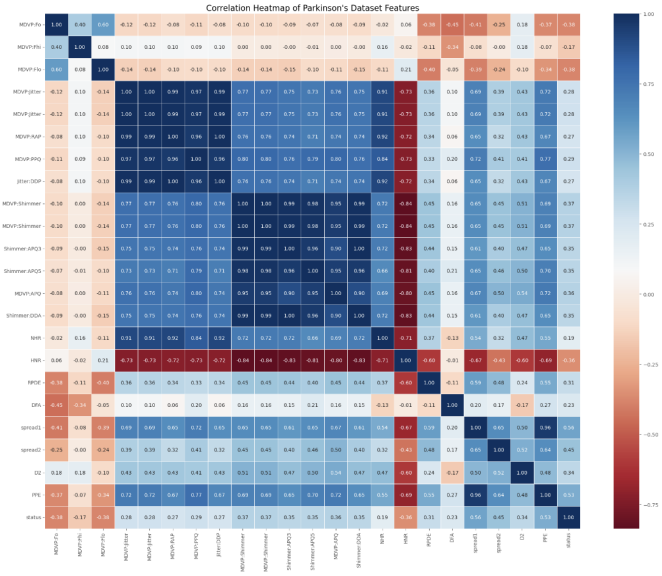

Fig. 4. Correlation HeatMap

process allowed each model to be specifically optimized for the Parkinson's dataset, maximizing diagnostic accuracy and minimizing overfitting risks. Optuna's integration ensured efficient model selection, significantly contributing to the framework's effectiveness in achieving early and accurate PD detection and establishing optimal classifier configurations.

D. Overview of Model Architectures and Implementation

1) Transformer: The Tabular Transformer model employs an advanced attention-based architecture tailored to structured data, such as the Parkinson's dataset. It integrates three core components: an embedding layer, multi-head self-attention layers, and feedforward layers, each essential for capturing complex feature dependencies and enhancing classification accuracy. The embedding layer transforms input vectors into a high-dimensional latent space, uniformly representing features of varying scales. At the model's core, multi-head self-attention layers generate sets of attention weights (heads) that analyze feature interactions, transforming inputs into Queries (Q), Keys (K), and Values (V) to compute attention scores. This mechanism captures intricate inter-feature relationships in Parkinson's data. Finally, the feedforward layer, a multi-layer perceptron with ReLU activation, refines high-level representations and decision boundaries, retaining relevant insights from attention layers. This architecture uniquely captures both simple and complex dependencies, making the Tabular Transformer a robust model for Parkinson's disease classification, outperforming traditional approaches in structured medical data applications.

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \qquad (1)$$

2) Convolutional Neural Network (CNN): The Convolutional Neural Network (CNN) model was tailored to analyze structured Parkinson's disease (PD) data, focusing on capturing intricate spatial relationships among features. Utilizing 1D convolutional layers, the model effectively processed feature-wise patterns in PD datasets. Each layer was followed by batch normalization for stable and faster training and dropout layers to reduce overfitting. Hyperparameter

optimization with Optuna fine-tuned the number of layers, filter sizes, and dropout rates. Global Average Pooling, rather than conventional max-pooling, was employed to retain more comprehensive spatial information. This customized CNN design, capable of learning both low- and high-level feature representations, proved effective in identifying complex PD-related patterns, enhancing classification accuracy beyond traditional methods.

$$h_{i,j} = \sigma\left(\sum_{k=0}^{K} w_k \cdot x_{i+k,j} + b\right) \tag{2}$$

3) Support Vector Machine: The Support Vector Machine (SVM) model in our study leverages a Radial Basis Function kernel, which is suitable for handling non-linear relationships in PD data. This kernel choice is pivotal because PD-related data features often exhibit complex, non-linear separability. In implementation, we utilized Optuna to optimize key hyperparameters such as the regularization parameter C and the kernel coefficient γ, enhancing the SVM's ability to find the optimal margin and avoid overfitting. The resulting model successfully handles the subtle distinctions in PD feature space, making it a valuable baseline for performance comparison.

4) Decision Tree: Decision trees were chosen for their interpretability and hierarchical structure, which is beneficial in understanding the importance of features in PD classification. Our implementation utilizes Gini impurity to guide the split at each node, constructing a tree that iteratively selects features and thresholds to minimize class entropy. To further refine the model, Optuna was employed to tune the maximum depth, minimum samples per leaf, and minimum samples required for a split, ensuring an optimal balance between model complexity and interpretability. The decision tree provides an easily interpretable model for PD classification with enhanced accuracy through careful parameter selection.

5) Random Forest: The Random Forest model builds upon Decision Trees by creating an ensemble of multiple trees using bootstrapped samples, where each tree is allowed to grow without pruning. This approach reduces the risk of overfitting and enhances model stability. The implementation included 200 decision trees, with hyperparameters like maximum tree depth and the number of features to consider for each split optimized through Optuna. The Random Forest model benefits from this ensembling approach, which averages predictions across the trees and ultimately boosts classification robustness for PD diagnosis.

6) Gradient Boosting Classifier: A gradient-boosting classifier was implemented to build an ensemble of weak learners (shallow trees) sequentially, where each learner attempts to correct the errors of its predecessors. Optuna was instrumental in tuning parameters such as learning rate, number of estimators, and tree depth. The incremental improvement of the ensemble through learning from previous errors enhances the model's ability to accurately classify complex PD patterns. By optimizing these hyperparameters, our GBC implementation achieves a strong balance between accuracy and overfitting control, making it a competitive model for PD classification.

$$F_m(x) = F_{m-1}(x) + \gamma_m h_m(x) \tag{3}$$

7) AdaBoost: The AdaBoost model was configured to adaptively focus on misclassified instances by adjusting their weights in subsequent rounds. Our implementation involved setting up a series of weak classifiers, with each successive classifier concentrating more on difficult-to-classify samples. Optuna was used to adjust the number of estimators and learning rate, ensuring the model converged to an optimal solution. This adaptive focus makes it particularly suited for handling challenging cases within the PD dataset, where early symptoms might be subtle and harder to distinguish.

8) XGBoost: The XGBoost model, known for its speed and accuracy, incorporates regularization to prevent overfitting. We implemented it with a parameter grid that includes max depth, learning rate, and subsample ratio. Optuna's hyperparameter tuning was essential here, allowing us to configure the tree's depth, learning rate, and regularization coefficients (L1 and L2) to improve performance specifically for PD data. By integrating L2 regularization, the model achieves a refined balance between complexity and bias, leading to improved generalization on PD classification tasks.

$$\mathcal{L} = \sum_{i=1}^{n} l(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k) \tag{4}$$

9) K-Nearest Neighbors (KNN): K-Nearest Neighbors (KNN) provides a simple yet effective non-parametric classification method. The KNN model implemented in this study uses Euclidean distance to determine the nearest neighbors, with hyperparameters such as the number of neighbors K and the weighting function optimized via Optuna. This approach allows for flexibility in PD classification, where instances are classified based on the majority class among their closest neighbors. Despite being a straightforward model, KNN's adaptability and interpretability make it valuable for our model comparison.

E. Model Evaluation Metrics

The evaluation of the proposed model was conducted using comprehensive metrics to ensure diagnostic reliability in classifying Parkinson's disease (PD).

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{5}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{6}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{7}$$

$$F1 = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{8}$$

Accuracy measures the proportion of correctly classified instances (true positives and true negatives) among all samples. Precision indicates the fraction of true positive predictions out of all positive predictions. Recall measures the model's ability to correctly identify all true positive cases. F1 Score provides a harmonic mean of precision and recall. These metrics collectively demonstrate the robustness and clinical applicability of the model in distinguishing PD-positive and healthy cases.

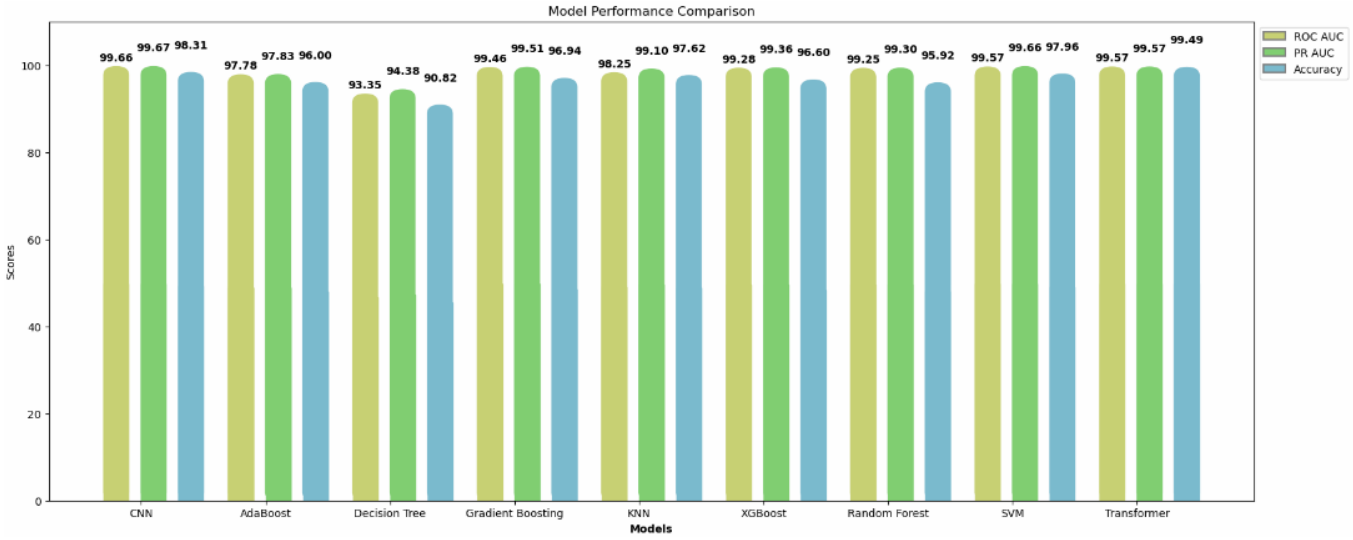

Fig. 5. ROC AUC, PR AUC, and Accuracy Analysis

IV. Results and Model Analysis

Table I provides a summary of the performance metrics for each model, highlighting the superior outcomes of deep learning models over traditional methods. The Transformer and CNN models demonstrated substantial classification efficacy, particularly the Transformer, which achieved the highest accuracy of 99%.

TABLE I
Performance Comparison of Different Models

Model	Accuracy (%)	F1 Score (%)	Precision (%)	Recall (%)
CNN	98.31	0.9831	0.98	0.98
AdaBoost	96.00	0.9600	0.96	0.96
Decision Tree	90.82	0.9082	0.91	0.91
Gradient Boosting	96.94	0.9694	0.97	0.97
KNN	97.62	0.9762	0.97	0.97
XGBoost	96.60	0.9660	0.97	0.97
Random Forest	95.92	0.9592	0.96	0.96
SVM	97.96	0.9796	0.98	0.98
Transformer	99.49	0.9949	0.99	0.99

A. Performance Analysis by Model

The performance analysis of various machine learning models revealed a clear superiority of deep learning techniques over traditional algorithms. The Transformer model emerged as the top performer, achieving an impressive accuracy of 99.49% with ROC AUC and PR AUC scores of 0.9957 and 0.9966, respectively. Its self-attention mechanism proved highly effective in capturing both short-range and long-range feature dependencies. The Convolutional Neural Network (CNN) followed closely, demonstrating strong capabilities in spatial feature extraction with an accuracy of 98.31% and AUC scores above 0.996. Among traditional classifiers, AdaBoost showed notable performance with 96% accuracy and AUC scores around 0.978, leveraging its ensemble-based architecture to handle dataset complexities. Other models, including Gradient Boosting, XGBoost, Random Forest, and SVM, achieved accuracies ranging from 90.82% to 97.96%, with Gradient Boosting standing out at 96.94% accuracy and 0.9946 ROC AUC. The superior performance of deep learning models, particularly the Transformer and CNN, underscores their potential in advancing machine learning applications in healthcare, especially for complex, multi-dimensional medical datasets. Their high recall and precision scores further emphasize their suitability for clinical settings where minimizing false negatives is crucial. This comprehensive evaluation highlights the evolving landscape of machine learning in medical diagnostics, with attention-based and convolutional architectures demonstrating particular promise in capturing intricate patterns in tabular medical data.

B. ROC AUC, Precision-Recall AUC, and Accuracy

Fig. 5 illustrates critical metrics such as ROC AUC, Precision-Recall AUC, and accuracy, which were analyzed to comprehensively assess diagnostic reliability. ROC AUC provided insights into the models' discrimination ability, while Precision-Recall AUC focused on performance for the positive class, which is crucial for minimizing false negatives in a diagnostic setting. Accuracy was considered in conjunction with these metrics to confirm the models' generalized effectiveness. This multidimensional evaluation approach highlights the clinical applicability of the deep learning models and underscores their capability for high-stakes medical data classification. Together, these metrics demonstrate that our tailored architectures especially Transformers and CNNs—are well-suited to the nuanced demands of medical diagnostics, ensuring reliable classification and minimizing the risks of false negatives.

Discussions

The Tabular Transformer represents a significant advancement, specifically addressing the challenges of structured data classification in Parkinson's disease (PD) detection. Distinguished from conventional deep learning approaches, this architecture leverages a specialized embedding layer that uniformly transforms raw features into a high-dimensional latent space, enabling robust feature representation independent of original scale or distribution. The model's multi-head self-attention mechanism dynamically captures complex inter-feature dependencies, a critical advantage in medical datasets where nuanced biomarker interactions are pivotal for accurate diagnostics. For instance, Srinivasan et al. (2024) [10] achieved 93.84% accuracy using Support Vector Machines (SVM) with Recursive Feature Elimination on voice-based features, while Singh et al. (2023) [11] reported enhanced accuracy using ensemble methods like Random

Forests and Gradient Boosting, though limited to a peak of 96%. Similarly, Camacho et al. (2023) [12] utilized Convolutional Neural Networks (CNNs) for MRI data and achieved notable accuracy, but these approaches often require complex preprocessing or are confined to specific data modalities like imaging.

By integrating advanced preprocessing techniques like SMOTE and Optuna, the Tabular Transformer effectively mitigates class imbalance and ensures consistent model performance. Empirical validation demonstrates superior classification accuracy of 99.49%. Unlike sequential Transformers and Vision Transformers designed for specific data modalities, this approach directly processes tabular data, efficiently capturing non-linear feature interactions and establishing a promising paradigm for structured medical data analysis.

The study also evaluates models like CNNs, SVM, Decision Tree, Random Forest, AdaBoost, Gradient Boosting, XGBoost, and KNN. CNNs excelled with 98.31% accuracy, while XGBoost handled complex patterns effectively. Ensemble methods like Random Forest and Gradient Boosting provided stability, and traditional models like SVM offered baseline comparability. This highlights the strengths of diverse approaches and the superiority of attention-based architectures for PD diagnostics.

Conclusion and Future Work

This study demonstrates new ground by adapting advanced deep learning models—Transformers and Convolutional Neural Networks (CNNs)—to structured tabular healthcare data, using Parkinson's disease as a case study. This methodology, integrating diverse models with advanced preprocessing and optimization techniques, enabled the development of a high-accuracy framework for PD diagnosis. This framework not only enhances early PD classification accuracy by addressing limitations in previous approaches but also paves the way for future advancements in PD diagnostics. By successfully applying these sophisticated models to tabular healthcare data, this research paves the way for enhanced diagnostic accuracy and efficiency in clinical settings, potentially revolutionizing the field of medical informatics.

Future work will focus on further enhancing these models by exploring transfer learning to enable even more robust generalization across diverse patient populations. Additionally, integrating multimodal data like clinical notes, genetic information, and imaging could refine diagnostic precision, facilitating a comprehensive view of the patient profile. Incorporating explainable AI techniques, like SHAP (Shapley Additive Explanations), will be essential to interpreting model outputs, providing transparency and insights into feature importance. This approach could be extended to other neurological conditions, thereby advancing predictive modeling in broader healthcare contexts. Furthermore, developing real-time, interpretable outputs from these models will ensure alignment with clinical needs, empowering healthcare professionals with actionable insights and fostering trust in AI-driven diagnostic tools.

References

[1] D. Aarsland *et al.*, "Neuropsychiatric and motor symptoms in parkinson's disease: Advances in diagnostic understanding," *Parkinsonism & Related Disorders*, vol. 101, pp. 1–8, 2022.

[2] M. J. Armstrong and M. S. Okun, "Diagnosis and treatment of parkinson disease: A review," *JAMA*, vol. 323, no. 6, pp. 548–560, 2020.

[3] R. B. Postuma *et al.*, "Early markers of parkinson disease progression: Clinical and biomarker insights," *Movement Disorders*, vol. 36, no. 3, pp. 531–541, 2021.

[4] B. P. Ghosh, T. Imam, N. Anjum, M. T. Mia, C. U. Siddiqua, K. S. Sharif, M. M. Khan, M. A. I. Mamun, and M. Z. Hossain, "Advancing chronic kidney disease prediction: Comparative analysis of machine learning algorithms and a hybrid model," *Journal of Computer Science and Technology Studies*, vol. 6, no. 3, pp. 15–21, 2024.

[5] E. Tolosa *et al.*, "Machine learning and parkinson's disease: Progress and challenges in early diagnosis," *Neuroinformatics*, vol. 21, no. 1, pp. 15–29, 2023.

[6] A. Kaur *et al.*, "Application of deep learning in neurological disorder diagnostics," *Artificial Intelligence in Medicine*, vol. 128, p. 102180, 2022.

[7] N. Musti *et al.*, "Enhancing early parkinson's disease detection through multimodal deep learning," *Nature Scientific Reports*, 2024.

[8] J. Zhang, "Mining imaging and clinical data with machine learning approaches for the diagnosis and early detection of parkinson's disease," *npj Parkinson's Disease*, vol. 8, no. 1, p. 13, 2022.

[9] T. N. Nguyen, M. M. Khan, M. Z. Hossain, K. S. Sharif, R. Das, and M. S. Haque, "Product Demand Forecasting For Inventory Management with Freight Transportation Services Index Using Advanced Neural Networks Algorithm," *American Journal of Computing and Engineering*, vol. 7, no. 4, pp. 50–58, 2024. [Online]. Available: https://ideas.repec.org/a/bfy/ojajce/v7y2024i4p50-58id2432.html

[10] M. Srinivasan *et al.*, "Detection of parkinson's disease using multiclass machine learning approach," *Scientific Reports*, 2024.

[11] P. Singh *et al.*, "Machine learning approaches for parkinson's disease classification: A review of traditional and modern techniques," *IEEE Access*, vol. 11, pp. 10 112–10 124, 2023.

[12] R. Camacho *et al.*, "Robust explainable deep-learning classifier for parkinson's disease using t1-weighted mri," *Nature Scientific Reports*, 2023.

[13] S. Thakur *et al.*, "Cnn model for accurate roi identification in datscan images for pd diagnosis," *IEEE Transactions on Medical Imaging*, vol. 41, no. 8, pp. 1972–1983, 2022.

[14] I. Khan *et al.*, "Transformers in medical data analysis: Application to disease classification," *Medical Informatics*, vol. 18, no. 4, pp. 200–213, 2021.

[15] H. Yoshida and T. Kameda, "Hyperparameter optimization for medical diagnostics using optuna: An ml and dl application," *Journal of Applied Artificial Intelligence*, vol. 35, no. 3, pp. 260–275, 2021.

[16] J. Thorp *et al.*, "Machine learning models for parkinson disease: Systematic review," *JMIR Medical Informatics*, 2024.

[17] M. Little, "Parkinsons," UCI Machine Learning Repository, 2007, [Online]. Available: https://doi.org/10.24432/C59C74.

[18] G. Aditi and P. Sushila, "Early detection of parkinson's disease using machine learning," *Procedia Computer Science*, vol. 217, pp. 612–621, 2023.

[19] A. H. Shahid and M. P. Singh, "A deep learning approach for prediction of parkinson's disease progression," *Biomedical Engineering Letters*, vol. 10, pp. 227–240, 2020.

[20] C. Marras, J. C. Beck, J. H. Bower, E. Roberts, B. Ritz, G. W. Ross, R. D. Abbott, R. Savica, S. K. Van Den Eeden, A. W. Willis *et al.*, "Prevalence of parkinson's disease across north america," *NPJ Parkinson's disease*, vol. 4, no. 1, pp. 1–7, 2018.