

Article

A Secure and Robust ML Framework for Sequence Classification and Adversarial Evaluation in a Bilateral Carpal Tunnel Syndrome Crossover Dataset

Pratik Pandurang Kharat ¹ , Sufian Al Majmaie ¹ , Ghazal Ghajari ¹ , Fathi Amsaad ¹ 
and Mohamed I. Ibrahim ^{2,3,*} 

¹ Department of Computer Science and Engineering, Wright State University, Dayton, OH 45435, USA; kharat.3@wright.edu (P.P.K.); al-majmaie.2@wright.edu (S.A.M.)

² Cyber Systems Engineering Department, Augusta University, Augusta, GA 30912, USA

³ Department of Electrical Engineering, Faculty of Engineering at Shoubra, Benha University, Cairo 11672, Egypt

* Correspondence: mibrahem@augusta.edu

Abstract

Bilateral idiopathic carpal tunnel syndrome (CTS) is a neuromuscular condition involving the compression of the median nerve at both wrists, leading to pain, neurological symptoms, and loss of function. This paper proposes a robust machine-learning framework for a randomized crossover clinical trial comparing two physiotherapeutic treatment regimens: stretching followed by myofascial mobilization (S/M) and the reverse sequence (M/S). Instead of making inferences about the superiority of one treatment over another, the treatment regimen serves as a structured analytical label for investigating predictive separability, feature representation, and model stability within a controlled experimental setting. The clinical dataset of 73 patients underwent rigorous preprocessing, including strength feature aggregation and principal component analysis (PCA). Various classifiers were evaluated, with CatBoost achieving an ROC-AUC of 0.985 and a test accuracy of 96.5%, while Random Forest demonstrated strong adversarial robustness with an adversarial accuracy of 96.83%. To assess robustness, clinically constrained perturbations were introduced into the PCA feature space, simulating realistic input variability. The findings indicate that ensemble learning algorithms can capture structured patterns in crossover clinical datasets and remain stable under low-magnitude adversarial perturbations. The study underscores the importance of robustness evaluation and interpretability when applying machine learning models to biomedical data, particularly in small and well-structured clinical cohorts.

Keywords: bilateral carpal tunnel syndrome; treatment sequence classification; crossover trial analysis; machine learning; adversarial robustness



Academic Editor: Anirban Bandyopadhyay

Received: 30 January 2026

Revised: 7 March 2026

Accepted: 16 March 2026

Published: 17 March 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

1.1. Background

Carpal Tunnel Syndrome (CTS) is among the most common types of entrapment neuropathy, affecting approximately 3–6% of the adult population [1]. CTS develops when the median nerve becomes compressed within the carpal tunnel, a narrow anatomical passage located on the palmar side of the wrist. Common clinical symptoms include numbness, tingling, and pain affecting the thumb, index, middle, and inner ring fingers, often accompanied by reduced grip strength and overall impairment of hand function [2]. While CTS

presents bilaterally in only a few instances, bilateral idiopathic CTS, i.e., spontaneous onset of symptoms in both hands without an external etiologic factor, is especially challenging in terms of diagnosis and management [3]. CTS affects women three times more than men [4], as women have smaller carpal canals than men, and women with CTS have significantly smaller canals than healthy women. Narrowing of the canal occurs on the asymptomatic side in patients who have unilateral symptoms, and is clearly noticeable, as well as more severe, and there is no relation between age and canal size [5].

Physiotherapy represents a first-line treatment option for mild to moderate CTS, especially in patients where surgery either is not recommended or is not desirable [6]. The more traditional conservative therapies are splinting of the wrist, stretching exercises, nerve gliding, and manual therapies like myofascial mobilization. The instrument-Assisted Soft Tissue Mobilization (IASTM), a type of myofascial release, has become increasingly popular due to its potential to decrease adhesions and facilitate neuromuscular reactivation [7]. The correct sequence in which such interventions are to be administered remains inadequately investigated; however, and anecdotal evidence and practitioner choice are more commonly used to make this decision than empirical evidence.

Increased developments in Artificial Intelligence (AI) and Machine Learning (ML) enable unparalleled applications of data-driven approaches in rehabilitation medicine [8,9]. ML algorithms, especially ensemble learning methods, show robust predictive abilities in a broad spectrum of medical applications, such as early disease identification, treatment response, and patient stratification [10,11]. With respect to CTS, ML can be used to recognize subtle patterns among a myriad of clinical and physiological measures, and potentially to predict which therapeutic sequence will result in the best outcome in a particular patient.

1.2. Problem Statement

Although the burden of CTS continues to grow with improved physiotherapeutic methods, the comparative effectiveness of different intervention sequences in bilateral idiopathic CTS remains insufficiently studied from a data-driven perspective. Existing crossover clinical trials provide structured, labeled datasets that are well-suited for evaluating whether machine learning models can identify separable patterns across treatment sequence groups. However, the application of robust, adversarially evaluated ML pipelines to such datasets has not been systematically explored. This study addresses this gap by applying secure and reliable ML methods to a randomized crossover CTS dataset, with the goal of evaluating predictive separability, model interpretability, and adversarial robustness in a controlled biomedical setting.

1.3. Objectives and Contributions

The overall aim of this study is to develop a secure and reliable machine-learning pipeline to evaluate the predictive separability of two physiotherapeutic treatment sequences—stretching followed by myofascial mobilization (S/M) or the reverse (M/S)—using a randomized crossover clinical dataset of female patients with bilateral idiopathic carpal tunnel syndrome (CTS), with particular emphasis on adversarial robustness and model interpretability.

Objectives:

1. Curate and preprocess a clinically annotated dataset including electrophysiological, sensory, strength, and diagnostic variables.
2. Engineer informative strength features by aggregating repeated grip/pinch measurements and applying PCA to reduce redundancy while preserving predictive variance.
3. Train and compare multiple supervised classifiers to predict the treatment sequence (s/m vs. m/s) and identify the most effective model.

4. Identify clinically relevant predictors using model-based feature importance and explainability analysis.
5. Evaluate clinical feasibility by validating generalization on a held-out test set using standard performance metrics.

Contributions:

- We formulate and study the *treatment-sequence prediction* problem for bilateral idiopathic CTS using a real randomized crossover clinical trial dataset.
- We propose an end-to-end ML pipeline that integrates preprocessing, PCA-based strength feature engineering, model selection, and explainability for structured biomedical sequence classification.
- We introduce a clinically constrained adversarial robustness evaluation in the PCA feature space to assess the security and reliability of the deployed prediction models.
- We provide an empirical comparison showing that ensemble methods (CatBoost and Random Forest) achieve high predictive performance while maintaining strong robustness under realistic perturbations.

1.4. Research Questions

1. Can machine learning algorithms identify statistically separable patterns between the S/M and M/S sequence groups in a randomized crossover CTS dataset?
2. What are the critical diagnostic and physiological characteristics used to evaluate treatment sequence?
3. How do various machine learning algorithms rank in predictive performance as applied to this task of classification?
4. What methodological insights can ML-based sequence classification provide for the design of future prospective studies in rehabilitation research?

1.5. Significance of the Study

This research presents an innovative strategy to the integration of machine learning analysis into personalized physiotherapeutic protocol design in CTS. Through its application of ML to a large clinical dataset of a crossover randomized controlled study [4], this research contributes to rehabilitation as well as health informatics, namely by:

1. Demonstrating the feasibility of applying robust ML pipelines to structured crossover clinical datasets in physiotherapy research.
2. Evaluating the stability and interpretability of ensemble classifiers under realistic adversarial perturbations in biomedical data.
3. Providing a methodological framework extensible to future prospective studies aiming to identify patient-level predictors of treatment response.

1.6. Justification of Treatment Sequence as Analytical Target

A central methodological consideration in this study concerns the use of the randomized treatment sequence label (S/M vs. M/S) as the classification target. Since sequence assignment was determined by randomization rather than by clinical judgment or patient characteristics, this choice warrants explicit justification.

The primary motivation for this analytical approach is methodological rather than directly clinical. The randomized crossover design of the source trial [4] provides a well-controlled, balanced, and unbiased labeling structure, precisely the conditions under which machine learning models can be rigorously evaluated for their ability to detect structured patterns in high-dimensional biomedical data. By using the sequence label as the target, this study investigates whether baseline clinical and physiological features, such as electrophysiological severity, sensory function, and hand strength, carry any systematic

association with the assigned sequence group, even under randomization. The detection of such patterns, if present, may reflect subtle pre-existing differences between groups that survived the randomization process, or may indicate features that are informative for future prospective stratification studies.

Importantly, this study makes no causal claim that predicting S/M versus M/S is equivalent to identifying the *optimal* treatment for a given patient. The models are not proposed as clinical decision tools ready for deployment. Rather, the framework serves three complementary purposes: (1) to evaluate the predictive separability of sequence groups using structured biomedical features; (2) to benchmark ensemble ML methods under realistic adversarial perturbations in a controlled clinical dataset; and (3) to establish a reproducible pipeline that can be extended to future studies where the target label is a clinically derived outcome measure, such as symptom severity change or functional recovery score.

This framing aligns with an emerging body of work that applies ML to randomized trial data not for direct clinical prediction, but for hypothesis generation, subgroup discovery, and methodological benchmarking [12]. The findings of this study should therefore be interpreted as exploratory and proof-of-concept in nature, with external validation and outcome-based redesign identified as necessary steps before any clinical translation.

1.7. Paper Organization

This paper is organized into the following main sections: Section 2 reviews CTS background, physiotherapy interventions, and related ML studies. Section 3 describes the dataset, preprocessing, feature engineering, models, and evaluation protocol. Section 4 reports experimental results and explainability findings. Section 5 presents the adversarial threat model and robustness analysis. Section 6 concludes the paper and outlines future directions.

2. Related Work and Literature Review

2.1. Overview of Carpal Tunnel Syndrome (CTS)

Carpal tunnel syndrome (CTS) was first described by Paget in 1854 and has since been recognized as a significant disabling condition, particularly in rheumatology and orthopedic hand surgery [13]. Among all median nerve compression disorders, CTS remains the most prevalent and extensively studied, accounting for approximately 90% of all entrapment neuropathies [14]. The condition is primarily attributed to increased pressure within the carpal tunnel—a confined anatomical space formed by the carpal bones and the transverse carpal ligament which leads to impaired median nerve function [15].

Clinically, CTS presents as compression of the median nerve at the wrist, resulting in symptoms such as pain, numbness, and tingling, primarily affecting the thumb, index finger, middle finger, and inner side of the ring finger [16,17].

The carpal tunnel is a narrow, rigid passageway located at the base of the hand, bordered laterally and inferiorly by the carpal bones and covered superiorly by the transverse carpal ligament—a dense, fibrous structure. This tunnel accommodates tendons that control finger movement, each encased in a lubricating sheath, along with the median nerve, which provides sensory and motor innervation to parts of the hand. In CTS, swelling of the tendons or surrounding soft tissues within this limited space compresses the median nerve, resulting in the characteristic clinical symptoms of pain, numbness, and tingling [18,19].

An epidemiological study conducted between September 1983 and July 1985 in Maastricht, Netherlands, and nearby villages assessed 501 individuals from the general population and identified 156 cases of CTS, comprising 131 women and 25 men [20]. Anatomical differences are believed to contribute to the higher prevalence of CTS among women;

specifically, females tend to have a smaller carpal tunnel, which increases the likelihood of median nerve compression. Research further suggests that women often exhibit a smaller carpal arch and reduced palmar bowing, both factors that are associated with a greater risk of developing CTS [21]. Figure 1 illustrates the anatomical structure of the carpal tunnel and how median nerve compression leads to CTS symptoms such as pain, numbness, and reduced hand function.

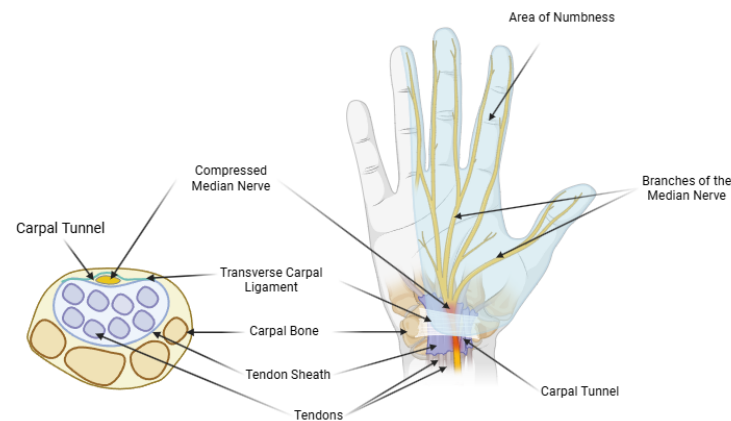


Figure 1. Anatomy of the carpal tunnel showing median nerve compression that leads to CTS symptoms.

2.2. Challenges in Bilateral Idiopathic CTS

Bilateral idiopathic Carpal Tunnel Syndrome (CTS) describes the compression of the median nerve at the wrist in both hands without an identifiable underlying cause. The term “idiopathic” indicates that the condition develops spontaneously or from unknown origins [22]. This variation in CTS is clinically significant because it tends to present with more intense symptoms and greater functional limitations compared to unilateral cases. The involvement of both hands substantially impacts daily living activities, often leading to considerable reductions in quality of life [3]. Diagnosing bilateral idiopathic CTS poses particular challenges, primarily due to the absence of a clear causative factor. Furthermore, the simultaneous involvement of both hands can obscure the severity of symptoms, potentially delaying medical intervention. Differentiating idiopathic CTS from other disorders with overlapping symptoms, such as peripheral neuropathies or systemic diseases, requires thorough clinical assessment and diagnostic precision [22,23]. Treatment strategies for bilateral idiopathic CTS also present unique obstacles. Conservative measures such as splinting or corticosteroid injections may prove less effective when both hands are simultaneously affected. Surgical approaches, including carpal tunnel release, demand careful scheduling and planning to avoid complete impairment of hand function during the recovery period. Prompt diagnosis and personalized treatment planning are essential for managing symptoms and maintaining hand functionality. Moreover, ongoing research focused on uncovering the causes and refining treatment strategies remains critical for improving clinical outcomes for individuals living with bilateral idiopathic CTS [22,24,25].

2.3. Physiotherapeutic Interventions

2.3.1. Stretching Techniques

Stretching exercises are commonly used in the management of Carpal Tunnel Syndrome (CTS) to reduce tension in the wrist and forearm muscles and to facilitate better mobility of the median nerve. Techniques such as nerve gliding and tendon gliding exercises are specifically designed to improve the excursion of the median nerve within the carpal tunnel. Evidence from a randomized controlled trial conducted by [26] demonstrated that patients with mild to moderate CTS who performed nerve gliding exercises

experienced significant reductions in symptom severity and improvements in hand function. These stretching methods help prevent the formation of nerve adhesions and reduce inflammation by enhancing fluid movement and minimizing mechanical irritation around the nerve structures [27].

2.3.2. Myofascial Mobilization

Myofascial mobilization is a therapeutic approach aimed at addressing soft tissue restrictions within the fascia and muscles of the forearm and wrist. The method involves the application of gentle, sustained pressure to the fascial tissues to release adhesions, alleviate tightness, and improve tissue elasticity. When applied to the management of Carpal Tunnel Syndrome (CTS), myofascial release helps to ease pressure on the median nerve by relaxing tightened muscles and surrounding fascial structures. Evidence from a study by Fernández de Las Peñas [28] demonstrated that manual therapy interventions, including myofascial techniques, were effective in significantly reducing pain and enhancing nerve conduction in patients diagnosed with CTS.

2.4. Machine Learning in Healthcare

Machine learning (ML) has emerged as a powerful tool for analyzing complex healthcare datasets, playing a critical role in disease diagnosis, treatment planning, and patient monitoring. ML models, for instance, have been successfully developed to predict patient outcomes during hospitalization and post-discharge, enabling clinicians to make better-informed medical decisions. In addition, ML drives the advancement of personalized medicine by allowing treatments to be customized based on individual patient data, significantly enhancing the effectiveness of therapeutic interventions [29].

2.4.1. Applications in Rehabilitation

In rehabilitation, ML enhances the assessment and monitoring of patient progress. Studies have demonstrated the use of ML in robotic rehabilitation to predict upper limb motor function recovery in stroke survivors [30]. Furthermore, ML algorithms have been applied to evaluate posture quality during rehabilitation exercises, providing objective assessments that can inform therapy adjustments [31]. These applications enable more precise and individualized rehabilitation strategies [32].

2.4.2. Advantages in Predictive Treatment Planning

- **Early Detection:** Machine learning models can help detect hidden cancers early. A recent study used ML to predict cancer in patients after venous thromboembolism, achieving 94% specificity and 93% negative predictive value, making it effective at ruling out cancer-free cases [33].
- **Improved Treatment Planning:** A personalized treatment plan can be generated using the patient's data and ML. This approach has been shown to enhance treatment outcomes, particularly in managing chronic conditions [34].
- **Data-Driven Decisions:** ML facilitates data-driven decision-making in healthcare, improving patient outcomes through predictive modeling [35,36].

2.5. Existing ML Systems for CTS

The existing system for managing bilateral idiopathic carpal tunnel syndrome relies on standardized, non-personalized physiotherapy protocols without predictive tools to guide the treatment sequence. Table 1 summarizes representative ML-based systems for CTS diagnosis and outcome prediction, highlighting the datasets, algorithms, reported performance, and key limitations that motivate the need for a more personalized sequence-prediction framework.

Table 1. Overview of existing machine learning approaches for carpal tunnel syndrome (CTS) diagnosis, including datasets, algorithms, performance metrics, strengths, and limitations.

Study	Algorithm	Dataset	Accuracy/Metric	Strengths	Drawbacks
Automatic Electro diagnosis of CTS Using ML [37]	SVM (95.13%), Random Forest, Logistic Regression, Decision Tree, Naïve Bayes	38 volunteers (28 CTS, 10 controls)	SVM: 95% (NCS), 89% (clinical)	Highly automated; electro diagnostic focus	Very small sample size, only signal-level data
Predicting Outcome After CTS Surgery [38]	Logistic Regression, Random Forest, Gradient Boosting Machine (best AUC 0.782)	1916 CTS post-surgery patients	Functional: 71.8%, Symptom: 75.9%	Real-world data, practical outcome prediction	Post-op only; lacks pre-surgery prediction modeling
ML-based CTS Severity Classification [39]	XGBoost (76.6%), Random Forest (75.6%), k-NN	1037 CTS hands	XGBoost: 76.6% external validation	Severity prediction with multi-grade classification	BMI and pain scores dominate model; moderate multiclass results
ML for Supporting CTS Surgery Decisions [40]	Chi-squared Automatic Interaction Detection (CHAID), ML classifiers	1916 CTS surgical registry	Functional: 71.8%, Symptomatic: 75.9%	Decision trees for easy clinical interpretation	Limited external validation beyond UK; fewer variables considered
Predicting CTS Diagnosis and Prognosis Using ML [41]	Bagging (Random Forest)	160 patients (80 CTS, 80 non-CTS)	Diagnosis: 95.5%, Prognosis (6 months): 91.2%	Comprehensive feature set; longitudinal tracking	Mixed population; possible symptom overlap bias
Deep Learning-Based CTS Diagnosis [42]	3 DL Models (CNN-based, SHAP explainability)	160 CTS patients	Diagnosis Accuracy: 96.9%, CSA prediction R^2 : 0.667	Multi-layer prediction; SHAP explainability; forecasting CSA changes	High reliance on ultrasound CSA; complex model stacking

2.6. Adversarial Risks in Clinical ML Models

Adversarial attacks pose a significant threat to the reliability and safety of AI models deployed in clinical settings. These attacks involve subtle, often imperceptible alterations to input data that can mislead AI systems into making incorrect predictions or classifications. In the context of healthcare, such vulnerabilities can have serious consequences, including misdiagnoses, inappropriate treatment plans, and threats to patient safety [43]. Finlayson [44] explores the motivations behind adversarial attacks in medical machine learning, noting that different stakeholders may exploit these vulnerabilities for financial gain or competitive advantage. The study emphasizes the importance of implementing proactive detection and mitigation strategies to safeguard clinical AI systems. In the field of medical imaging, Ma [45] reviews recent advancements in adversarial attack methodologies targeting deep neural networks and discusses defense techniques designed to counter these threats. The research underscores the need for developing robust AI models that can maintain diagnostic accuracy even when faced with adversarial manipulations. Collectively, the existing literature highlights the critical need to secure AI systems in healthcare against adversarial risks. Sustained collaboration among clinicians, AI developers, and cybersecurity experts is essential to design and implement effective defense strategies, ensuring the safe and reliable integration of AI technologies into clinical practice [46].

3. Methodology

3.1. Dataset Description

This study utilizes a comprehensive dataset collected as part of a randomized, double-blind, crossover-controlled clinical trial that explored two physiotherapeutic treatment sequences, Stretching and Myofascial Mobilization (IASTM), in patients diagnosed with bilateral idiopathic carpal tunnel syndrome (CTS). The trial enrolled 73 participants from Londrina, Brazil, including 43 individuals who underwent surgical decompression and 30 who received conservative treatment. Each participant experienced both therapeutic interventions in alternating sequences across four treatment phases and a baseline phase, spanning T0 to T4 [4].

The dataset is structured into six principal blocks: experimental design specifications (Block 0), sociodemographic information (Block 1), electrophysiological diagnostic criteria (Block 2), objective strength evaluations (Block 3), sensory test results (Block 4), and patient-reported subjective outcomes (Block 5). Data collection involved repeated clinical assessments and self-report instruments, systematically tracked throughout the crossover study phases to support both within-subject and between-subject analyses.

Objective strength measurements included grip and pinch force assessments conducted with hydraulic devices and repeated three times per hand. These repeated values were averaged to minimize measurement variability and enhance reliability. To further refine the strength data, Principal Component Analysis (PCA) was performed on 12 grip and pinch strength variables. This dimensionality reduction produced two principal components, Strength_PCA1 and Strength_PCA2. Strength_PCA1 primarily captured overall hand strength, while Strength_PCA2 reflected asymmetry and variations between grip and pinch strength across hands. The severity of electrophysiological impairment was classified using established diagnostic frameworks, including the Stevens, Padua, Bland, and Hiob-DB systems, providing a detailed and standardized evaluation of CTS severity for each hand. In addition, sensory deficits were assessed through the Semmes–Weinstein monofilament test and the Two-Point Discrimination Test, offering a comprehensive view of sensory function among participants.

3.2. Feature Engineering

Feature engineering was conducted to ensure that the dataset was both analytically robust and consistent with domain-specific knowledge [47] about carpal tunnel syndrome. This process involved multiple phases, including data transformation, quality assurance, and dimensionality reduction. Initially, placeholder values coded as 999 used to denote missing or inapplicable responses were systematically removed to prevent potential bias during model training and statistical evaluation. Furthermore, as described in Section 3.1, male and left-hand dominant participants were excluded, resulting in a final cohort of female, right-hand-dominant patients. Next, the repeated grip and pinch strength measurements were averaged for each hand and strength type to create new features. These aggregated features were then subjected to Principal Component Analysis (PCA). PCA was employed to reduce the dimensionality of the highly correlated strength measurements into more manageable and interpretable components. The resulting components Strength_PCA1 and Strength_PCA2 captured over 90% of the original variance and served as key features for subsequent machine learning models, enhancing their predictive performance.

The scatter plot in Figure 2, titled “PCA on Grip and Pinch Strength,” illustrates the application of Principal Component Analysis (PCA) to the raw grip and pinch strength measurements collected in the clinical dataset. PCA is a statistical technique used for dimensionality reduction, enabling high-dimensional data to be represented in a lower-dimensional space while preserving as much variance as possible [48,49]. This transfor-

mation supports effective visualization, noise reduction, and preparation for downstream machine learning tasks.

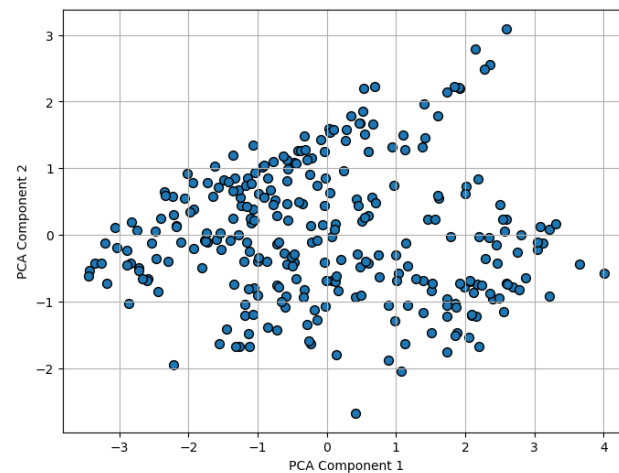


Figure 2. PCA scatter plot of grip and pinch strength measurements projected onto the first two principal components (Strength_PCA1 and Strength_PCA2).

In this analysis, multiple grip and pinch strength measurements were first standardized and mean-centered. PCA was then applied to identify the directions (principal components) along which the data exhibited the greatest variation. The dataset was projected onto the first two principal components (PC1 and PC2), which captured the most substantial variance in strength-related metrics. Each point in the scatter plot corresponds to an individual patient observation projected into this new coordinate system.

Strength Aggregation and PCA

Let $X \in \mathbb{R}^{n \times p}$ represent the matrix of original grip and pinch strength data, where n is the number of samples and p is the number of input features.

1. **Standardization (Mean Centering):**

Each column (feature) is mean-centered to ensure zero mean:

$$X_{\text{centered}} = X - \bar{X}$$

2. **Covariance Matrix Computation:**

$$\Sigma = \frac{1}{n-1} X_{\text{centered}}^T X_{\text{centered}}$$

3. **Eigen Decomposition:**

$$\Sigma W = W \Lambda$$

where W is the matrix of eigenvectors (principal components), and Λ is the diagonal matrix of eigenvalues (explained variance).

4. **Data Projection onto Principal Components:**

$$Z = X_{\text{centered}} W$$

The resulting matrix Z contains the transformed data, where each column corresponds to a principal component. In this study, the first two components, Z_1 and Z_2 , represent **Strength_PCA1** and **Strength_PCA2**, and were used to produce the PCA scatter plot:

$$Z^{(2)} = [\text{Strength_PCA1}, \text{Strength_PCA2}]$$

The scatter plot visually confirms that most data points cluster along these two components, validating their effectiveness in capturing the primary modes of variation in grip and pinch strength.

3.3. Data Preprocessing

To improve and enhance the data quality, consistency, and relevance for the prediction model, preprocessing is done on the clinical dataset using a structured series of steps applied to the dataset. The invalid data entries or incomplete values, known as missing values, are denoted by 999 and NaN. The rows containing these values were removed to keep the integrity of the datasets. Outlier detection procedures were applied to exclude entries that were not considered to be within the study inclusion criteria. As detailed in Section 3.1, male and left-hand dominant participants were excluded prior to analysis.

In the feature engineering phase, the average of the repeated grip and pinch strength readings is used to create more stable and representative metrics. To reduce the dimensionality and summarize the strength-related features, Principal Component Analysis (PCA) is applied to those features and converted into two primary components: Strength_PCA1 and Strength_PCA2. These components captured the most significant variation in the strength data and reduced the noise in the data. Feature selection is made on the dataset, keeping the features with high variance or low redundancy to optimize model performance as identified through correlation analysis. This selection process was further clarified using feature importance scores obtained from preliminary baseline models while keeping only the most impactful predictors. The cleaned and engineered dataset was partitioned into training and testing subsets, with 80% allocated for training and 20% reserved for testing. This data separation was essential to ensure that the model could generalize effectively and achieve greater predictive accuracy. Overall, the preprocessing pipeline established a strong foundation for the subsequent modeling and analysis phases.

3.4. Model Architecture

3.4.1. Model Selection and Design

The proposed machine learning pipeline was developed to classify and evaluate the separability of physiotherapy treatment sequences in a randomized crossover dataset of bilateral CTS patients. It integrates preprocessing steps, dimensionality reduction techniques, and ensemble-based classification models. The system utilizes a combination of clinical and strength-related features, including those derived from Principal Component Analysis (PCA), to improve both model performance and interpretability. The workflow incorporates multiple classifiers, cross-validation procedures, and adversarial testing to ensure model robustness and reliability under realistic clinical conditions. Figure 3 presents the end-to-end workflow of the proposed ML-assisted framework, including preprocessing, PCA-based feature engineering, model training/validation, and adversarial robustness assessment.

1. **Data Acquisition:** Collection of the clinical dataset related to bilateral idiopathic carpal tunnel syndrome (CTS).
2. **Data Preprocessing:**
 - Eliminate missing values and invalid entries from the dataset.
 - Filter the dataset to include only right-hand dominant female subjects.
 - Average repeated grip and pinch strength measurements.
3. **Feature Engineering and Selection:**
 - Apply Principal Component Analysis (PCA) to generate Strength_PCA1 and Strength_PCA2.
 - Remove low-variance and highly correlated features.
 - Retain top features based on feature importance scores from baseline models.

4. **Data Partitioning:**
 - Split the dataset into 80% training and 20% testing sets using stratified sampling.
5. **Model Training and Evaluation:**
 - Train models including Random Forest, CatBoost, XGBoost, SVM, Decision Tree, and AdaBoost.
 - Perform 10-fold cross-validation on each model.
 - Evaluate using Accuracy, Precision, Recall, F1-score, ROC-AUC, and Confusion Matrix.
6. **Adversarial Robustness Testing:**
 - Apply controlled perturbations to PCA features (Strength_PCA1 and Strength_PCA2).
 - Record Adversarial Accuracy and Attack Success Rate for each model.
7. **Final Model Selection and Interpretation:**
 - Select the model with the best robustness and predictive performance.
 - Use feature importance analysis to interpret clinically relevant predictors.

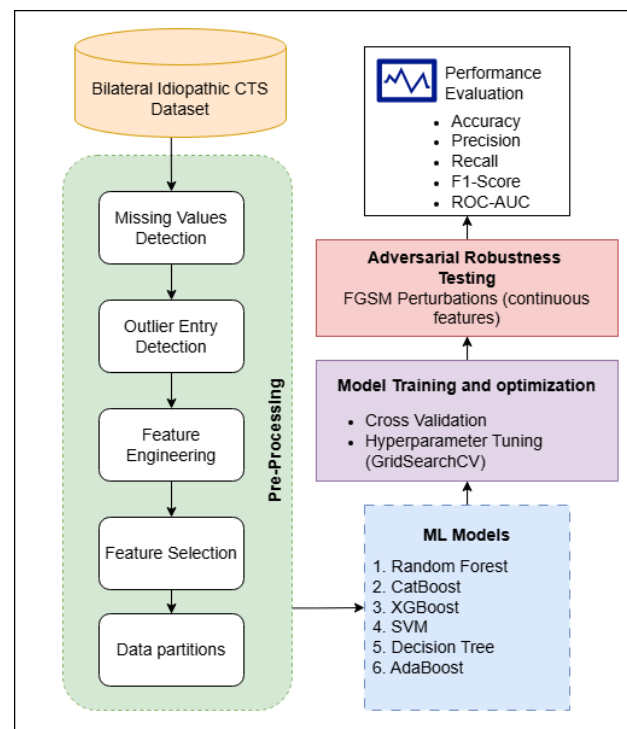


Figure 3. Machine learning pipeline showing data preprocessing, PCA-based feature engineering, model training, evaluation, and adversarial robustness testing.

3.4.2. Hyperparameter Tuning (e.g., GridSearchCV)

Hyperparameter tuning was performed using the GridSearchCV utility from the scikit-learn library to optimize the performance of each machine learning model. Hyperparameters—unlike model parameters learned during training, are external configurations (e.g., number of trees, learning rate, maximum tree depth) that directly influence how the algorithm learns from data [50].

The GridSearchCV function systematically evaluates a predefined set of hyperparameter combinations by performing an exhaustive search over a specified grid [51]. For each combination, the model is trained and validated using 10-fold cross-validation, ensuring that the evaluation is robust and not dependent on a particular train-test split. The cross-validated scoring metric (e.g., accuracy or AUC) is used to determine the best-performing parameter configuration.

This tuning process was applied consistently across models:

For Random Forest, parameters like `n_estimators`, `max_depth`, and `min_samples_split` were optimized.

For CatBoost, parameters such as `iterations`, `depth`, and `learning_rate` were fine-tuned.

Similar grid-based tuning was applied for XGBoost, SVM, Decision Tree, and AdaBoost, each with model-specific hyperparameters.

This approach helps prevent both overfitting and underfitting, improves generalization on unseen data, and ensures fair comparison between models trained under optimal configurations.

3.4.3. Cross-Validation Strategy

In this study, k-fold cross-validation was employed as a model evaluation strategy to ensure the generalizability and reliability of the machine learning classifiers predicting physiotherapy treatment sequences in bilateral CTS. Specifically, a 10-fold cross-validation procedure was implemented using the `GridSearchCV` function from `scikit-learn` [51], which performs both model validation and hyperparameter tuning simultaneously.

The process involves dividing the entire training dataset into 10 approximately equal-sized folds. For each iteration, one fold is held out as the validation set while the model is trained on the remaining nine folds. This process is repeated 10 times, such that each fold serves as a validation set exactly once. The performance metrics (e.g., accuracy, precision, F1-score, ROC-AUC) are averaged across all folds to provide a stable estimate of the model's effectiveness on unseen data. The use of stratified sampling ensures that the class distribution remains consistent across all folds, which is particularly important given the binary classification nature of the `B0_SEQUENCE` target variable in this study.

This cross-validation framework was consistently applied across all six classifiers (Random Forest, CatBoost, XGBoost, SVM, Decision Tree, AdaBoost), ensuring fair comparison and robust model selection [52].

3.5. Treatment Sequence Labeling

This study was designed as a prospective, randomized, double-blind crossover trial comparing two physiotherapeutic interventions for the treatment of bilateral idiopathic Carpal Tunnel Syndrome (CTS) [4]: Static Active Stretching (S) and Myofascial Mobilization using IASTM (M). A total of 73 participants were recruited, including 43 surgical patients (post-carpal tunnel release) and 30 clinical patients managed conservatively without surgery; the randomized crossover design allowed each participant to undergo both therapeutic interventions in different sequences, facilitating within-subject comparisons and minimizing variability between individuals. Participants were randomly assigned to one of two groups: the S/M group, which received stretching followed by myofascial mobilization, or the M/S group, which received myofascial mobilization followed by stretching. These assignments are labeled in the dataset under the variable `B0_SEQUENCE`, with values 1 for S/M and 2 for M/S. Each participant completed a first 4-week treatment phase, followed by a 1-week washout period, and then a second 4-week phase with the alternate therapy. Supervised therapy sessions were conducted twice weekly, including 15 structured exercises targeting the neck, shoulders, wrists, and fingers.

To ensure methodological rigor, randomization was implemented using opaque sealed envelopes, managed by an independent professional not involved in the study. A double-blind protocol was maintained throughout; neither the therapists nor the participants were aware of the treatment order. Treatments were administered symmetrically on both upper limbs, with stretching involving pain-free holds at maximum range for 40 s, and myofascial

mobilization performed using metal instruments (crochets) following a structured protocol that included digital palpation, tool palpation, and mobilization phases.

Treatment-related data were recorded in the dataset under Block 0 (B0_) variables, including B0_TREATMENT (0 = none, 1 = stretching, 2 = myofascial), B0_PERIOD (time points T0 to T4), and B0_SEQUENCE (indicating treatment order). This labeling framework enables clear tracking of the initial treatment, crossover point, and sequence-specific outcomes. The study title is: “Crossover Study of Two Combined Physiotherapeutic Treatment Methods on Chirurgical and Clinical Patients with Idiopathic bilateral carpal tunnel syndrome (CTS).”

Dataset Labeling for S/M vs. M/S

The recruitment and screening process initially involved the assessment of 252 post-surgical patients. Of these, 207 patients were excluded for not meeting the eligibility criteria, resulting in 45 patients who satisfied the inclusion requirements. However, following consent, two participants were excluded, one due to loss of contact and another due to withdrawal from the study, leading to a final cohort of 43 randomized participants.

At baseline (T0), participants were randomly assigned into two groups: the S/M group (n = 21), who received Stretching first followed by Myofascial Mobilization, and the M/S group (n = 22), who received Myofascial Mobilization first followed by Stretching. Each intervention phase lasted for four weeks, after which participants underwent a one-week washout period before crossing over to the alternate therapy for another four weeks of treatment. During follow-up, no participants were lost at the three-month evaluation (T3). However, by the six-month evaluation (T4), two participants were lost from the S/M group and one participant from the M/S group. Consequently, the final analysis included 19 participants in the S/M group and 21 participants in the M/S group. Figure 4 depicts the CONSORT-style enrollment and allocation process for the surgical (post-decompression) cohort used in the crossover trial, clarifying exclusions, randomization, and follow-up counts.

The recruitment and screening process for the clinical cohort involved the assessment of 394 patients. Out of these, 360 patients were excluded for not meeting the eligibility criteria, resulting in 34 patients who satisfied the inclusion requirements. Subsequently, four participants were excluded after consent, one due to loss of contact and three who voluntarily withdrew, leading to a final cohort of 30 randomized participants.

At baseline (T0), participants were evenly assigned into two groups: the S/M group (n = 15), which received Stretching first, followed by Myofascial Mobilization, and the M/S group (n = 15), which received Myofascial Mobilization first followed by Stretching. Each intervention phase lasted for four weeks, followed by a one-week washout period, after which participants crossed over to the alternate therapy for an additional four weeks. No participants were lost during the three-month follow-up (T3). By the six-month evaluation (T4), two participants were lost from the M/S group, while the S/M group experienced no loss. Thus, the final analysis included 15 participants in the S/M group and 13 participants in the M/S group. Figure 5 shows the corresponding enrollment and allocation flow for the clinical (non-surgical) cohort, outlining eligibility screening, randomization, and participant retention across follow-up periods.

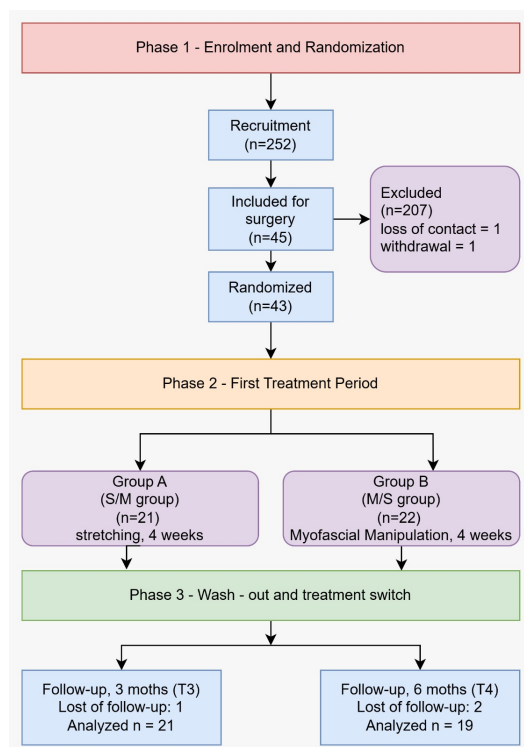


Figure 4. Flowchart of participant recruitment, screening, and allocation for the surgical cohort in the crossover CTS trial [4].

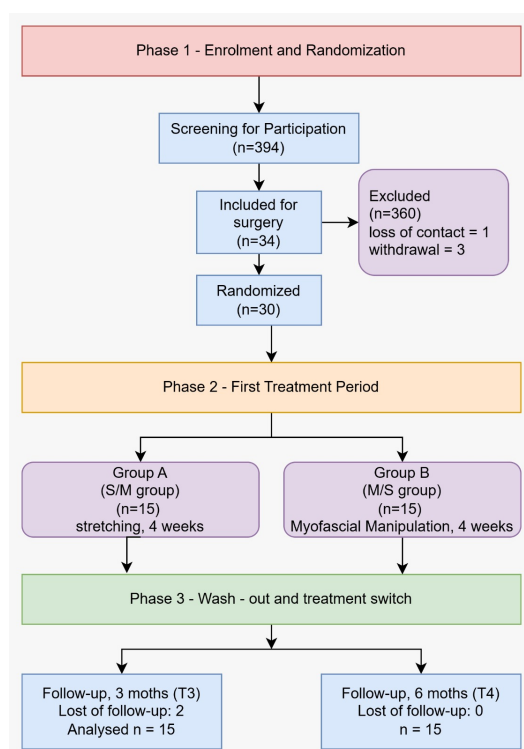


Figure 5. Flowchart of participant screening, eligibility assessment, and allocation for the clinical (non-surgical) CTS cohort [4].

4. Experimental Evaluation

4.1. Evaluation Metrics

4.1.1. Accuracy

Accuracy measures the proportion of correct predictions among all predictions.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

where

- TP—True Positives
- TN—True Negatives
- FP—False Positives
- FN—False Negatives

A high accuracy means the model correctly classifies most examples. However, in imbalanced datasets, accuracy alone may be misleading.

4.1.2. Precision

Precision (also called Positive Predictive Value) indicates how many predicted positives are actually positive.

$$\text{Precision} = \frac{TP}{TP + FP} \quad (2)$$

High precision indicates few false positives. This is crucial in medical diagnosis, where unnecessary treatment should be minimized.

4.1.3. Recall

Recall (also called Sensitivity or True Positive Rate) measures how many actual positives were correctly predicted.

$$\text{Recall} = \frac{TP}{TP + FN} \quad (3)$$

High recall is vital when missing a positive case has serious consequences (e.g., not diagnosing a patient who needs treatment).

4.1.4. F1-Score

The F1-score is the harmonic mean of precision and recall and provides a balance between the two.

$$\text{F1-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (4)$$

F1-score is especially useful when there is class imbalance and both false positives and false negatives are costly.

4.1.5. ROC-AUC Curve

The ROC curve plots the True Positive Rate (Recall) against the False Positive Rate (FPR) at various threshold settings. The AUC (Area Under Curve) summarizes the ROC curve into a single value.

$$\text{FPR} = \frac{FP}{FP + TN} \quad (5)$$

AUC is typically calculated via trapezoidal integration of the ROC curve. AUC ranges from 0.5 (random) to 1.0 (perfect). AUC is robust to class imbalance and gives a general sense of model ranking quality.

4.2. Model Performance

4.2.1. Random Forest Results

The Random Forest classifier demonstrated strong performance in classifying physiotherapy treatment sequences for patients with bilateral carpal tunnel syndrome (CTS). Model tuning was performed using a grid of hyperparameters, specifically: `max_depth = 10`, `max_features = 'sqrt'`, `min_samples_split = 5`, and `n_estimators = 200`. This configuration was selected to achieve an optimal balance between model complexity and generalizability, thereby reducing the risk of overfitting. Under 10-fold cross-validation, Random Forest yielded an 84.6% ($\pm 5.6\%$) average accuracy, 85.2% precision, 85.2% recall, and an F1-score of 84.6%, which verifies a high sensitivity–specificity balance. The ROC-AUC measure was highly significant at 92.4%, which demonstrates high discriminatory potential between treatment classes S/M and M/S. This indicates the model can reliably distinguish between S/M and M/S sequence labels within this crossover dataset. Figure 6 reports the confusion matrix of the Random Forest model on the test set, highlighting the distribution of correct classifications and the remaining misclassification patterns between S/M and M/S.

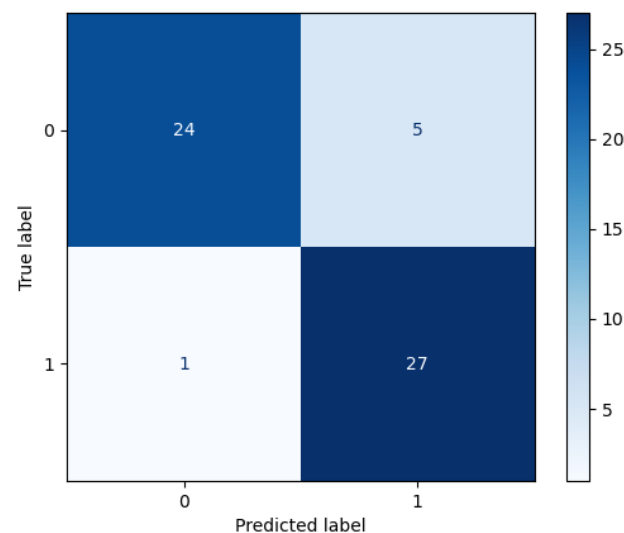


Figure 6. Confusion Matrix of the Random Forest Classifier. Matrix showing the distribution of correct and incorrect predictions for the classification of treatment sequence labels (S/M and M/S) on the test dataset.

On the hold-out test set, the model classified 89.5% of instances correctly, with 0.96 precision on 0 (S/M) and 0.84 on 1 (M/S), and 0.96 recall on M/S, which is particularly significant given clinical settings where missing a correct sequence would result in suboptimal treatment. The confusion matrix also indicates that, of 57 instances, only 6 are misclassified, 5 are false positives, and 1 is a false negative. This suggests the model is somewhat conservative toward the prediction of M/S treatment but infrequently overlooks actual positives. Figure 7 illustrates the ROC curve of the Random Forest classifier, demonstrating strong separability between the two treatment-sequence classes via a high AUC score.

An AUC of 0.98 was achieved for the test set using the ROC curve, which visually verified high separability between classes. The curve for testing vs. training accuracy also showed that while the accuracy for training ranged from 98.5%, the accuracy for testing settled between 91 and 93%, indicating that minor overfitting was overcome using parameter adjustment. These findings cumulatively show that the Random Forest model is a reliable and interpretable option for clinical sequence classification, offering high

performance and transparency of the algorithm, which is particularly important for clinical decision support systems operating in the real world.

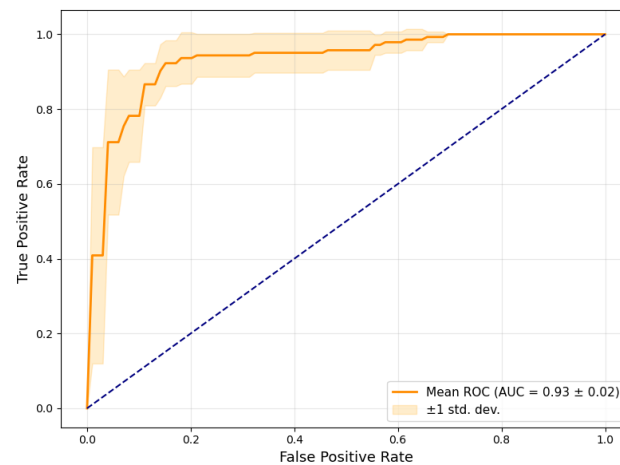


Figure 7. ROC curve of the Random Forest model showing classification performance for the treatment sequences.

4.2.2. CatBoost Results

The CatBoost classifier showed outstanding performance in predicting the best treatment sequence for bilateral carpal tunnel syndrome (CTS) patients. On the basis of stratified 10-fold cross-validation, the model reported a mean accuracy of 92.0% ($\pm 2.7\%$), precision of 91.7%, recall of 93.0%, F1-score of 92.2%, and an impressive mean ROC-AUC of 98.5% ($\pm 0.6\%$). These performances show that the model achieved an excellent trade-off between sensitivity and specificity, which is very important in clinical decision-making scenarios. Figure 8 presents the CatBoost confusion matrix on the hold-out test set, emphasizing the low false-negative rate that is especially desirable for safe clinical decision support.

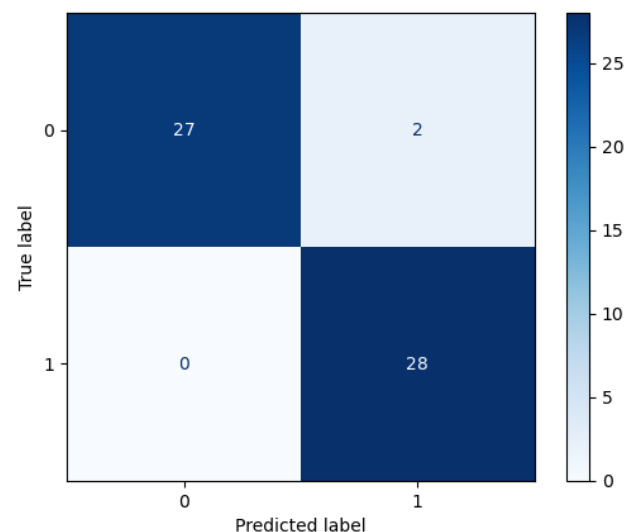


Figure 8. Confusion matrix of the CatBoost classifier showing predictions for the S/M and M/S treatment sequences.

On application to the hold-out test dataset, the accuracy on the test dataset was 96.49%, further solidifying its generalizability and robustness. The confusion matrix reflects this: out of 57 predictions, 27 true negatives and 28 true positives were both correctly predicted, and only 2 false positives and no false negatives. For example, avoiding false negatives ensures no patient is assigned an ineffective therapy sequence. This is significant, especially

since it shows that no true positive (M/S treatment group) instances were missed, with high clinical sensitivity. Figure 9 shows the ROC curve for CatBoost, where the near-perfect AUC indicates a highly accurate ranking of patients across treatment-sequence classes.

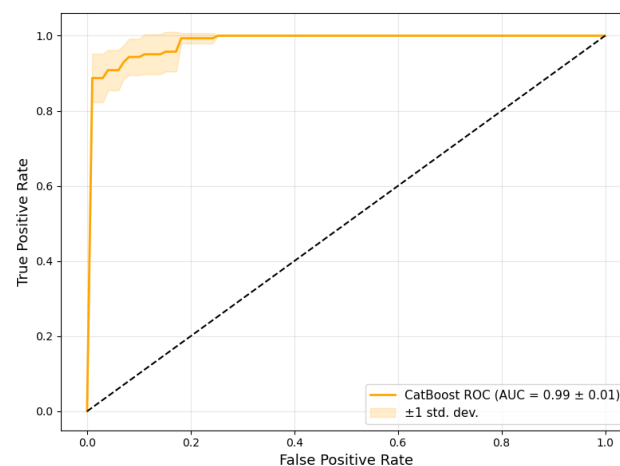


Figure 9. ROC curve of the CatBoost model showing classification performance for the treatment sequences.

In Figure 10, the curve from the ROC analysis for the test set demonstrates nearly perfect discrimination between classes, and an AUC of 1.00 indicates that the predictor distinguished between S/M and M/S sequences without error at all thresholds. This kind of performance would be particularly important for use within a clinical environment where suboptimal treatment allocation due to false negatives is undesirable.

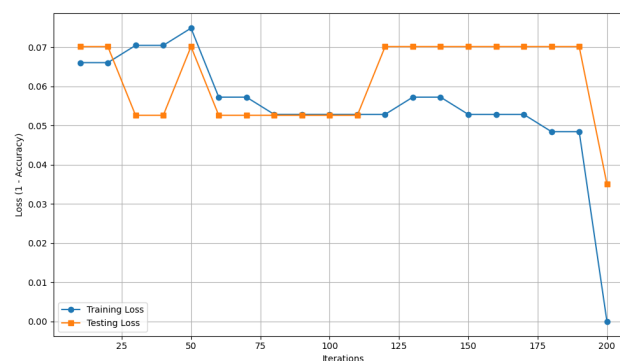


Figure 10. Training and testing performance curve of the CatBoost model across iterations.

Training dynamics were further analyzed using a curve comparing training and testing loss, expressed as $(1 - \text{accuracy})$. The curve remained relatively flat across a wide range of iterations, indicating that the model avoided significant overfitting or underfitting. The minimal gap observed between training and testing loss suggested that the model maintained low variance without sacrificing generalization capabilities. The optimal hyperparameters identified through grid search were: depth = 6, iterations = 200, learning_rate = 0.01, and l2_leaf_reg = 1. This configuration successfully balanced model complexity and regularization, resulting in high predictive accuracy and minimal variance. Taken together, these results demonstrate that CatBoost demonstrated strong classification performance for the treatment-sequence labels within this crossover dataset. for bilateral CTS. The model exhibits strong discriminative performance, reliability, and resilience to overfitting, making it a strong candidate for further evaluation in prospective clinical settings.

4.3. Comparison of Models

Table 2 summarizes the comparative performance of six machine learning models applied to classify physiotherapy treatment sequences in patients with bilateral carpal tunnel syndrome (CTS).

Table 2. Performance comparison of machine learning models for predicting physiotherapy treatment sequences in the CTS dataset using metrics such as ROC-AUC, accuracy, precision, recall, and F1-score.

Model	Best ROC-AUC	CV Accuracy	Test Accuracy	Test Precision	Test Recall	Test F1-Score
CatBoost	0.985	0.920 ± 0.027	0.965	1.00/0.93	0.93/1.00	0.96/0.97
Random Forest	0.924	0.846 ± 0.056	0.895	0.96/0.84	0.83/0.96	0.89/0.90
XGBoost	0.903	0.828 ± 0.045	0.860	1.00/0.78	0.72/1.00	0.84/0.88
AdaBoost	0.941	0.849 ± 0.088	0.860	0.92/0.81	0.79/0.93	0.85/0.87
Decision Tree	0.917	0.827 ± 0.099	0.789	0.79/0.79	0.79/0.79	0.79/0.79
SVM	0.853	0.801 ± 0.095	0.825	0.83/0.82	0.83/0.82	0.83/0.82
CatBoost (Without B0_CARRYOVER)	0.902	0.767 ± 0.061	0.842	0.92/0.79	0.76/0.93	0.83/0.85

Among all models evaluated, shown in Figure 11, CatBoost demonstrated the highest performance, achieving an ROC-AUC of 0.985, a cross-validation (CV) accuracy of 92.0% ($\pm 2.7\%$), and a test accuracy of 96.5%. It also attained excellent precision (1.00/0.93) and recall (0.93/1.00) across both classes, resulting in the highest F1-scores (0.96/0.97). Random Forest closely followed, with a test accuracy of 89.5%, a balanced F1-score of 0.89/0.90, and an ROC-AUC of 0.924. Although XGBoost exhibited slightly lower CV performance ($82.8 \pm 4.5\%$) and an ROC-AUC of 0.903, it maintained strong recall in one class (1.00) but showed imbalance in the other. AdaBoost performed competitively as well, achieving a test accuracy of 86.0% and an ROC-AUC of 0.941, offering a good trade-off between model complexity and generalization.

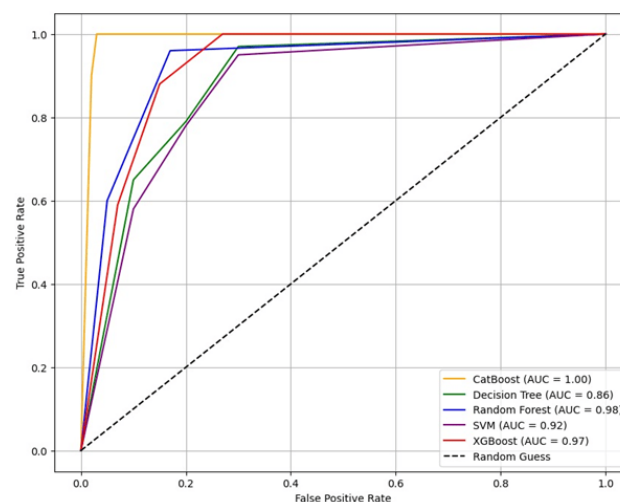


Figure 11. Comparison of ROC Curves for Ensemble Models. ROC curves comparing the classification performance of all models on the CTS dataset.

In contrast, the Decision Tree model exhibited greater variability and lower performance across all metrics, with a test accuracy of 78.9% and an F1-score of 0.79/0.79, reflecting its propensity to overfit. Finally, the SVM model achieved a respectable test accuracy of 82.5%, but its lower ROC-AUC of 0.853 and narrower F1-score margins (0.83/0.82) indicated limited robustness compared to the ensemble-based approaches. Overall, these findings underscore the superiority of tree-based ensemble methods, particularly CatBoost and Random Forest for this clinical prediction task.

4.4. Feature Importance and Explainability

A critical aspect of clinical machine learning models is explainability, which ensures transparency, builds trust, and improves the interpretability of predictive results. In this study, feature importance analysis was performed to determine which input variables most significantly influenced the model's decisions. Since the primary models used (Random Forest, XGBoost, and CatBoost) are tree-based, they inherently support impurity-based and permutation-based feature importance techniques. These methods enable ranking features based on how much they reduce node impurity across the ensemble or by evaluating how model accuracy changes when features are randomly permuted.

For dimensionality reduction, Principal Component Analysis (PCA) was applied to the strength-related features to derive Strength_PCA1 and Strength_PCA2. These PCA-transformed features captured the principal variance directions in the grip and pinch strength data, facilitating more stable visualization and explanation during classification and adversarial evaluation.

Understanding feature importance is crucial not only for enhancing model transparency and reliability but also for fostering clinical trust, detecting biases, and informing clinical interpretation. As emphasized by Molnar [53], interpretable machine learning is essential for building trust in AI applications within healthcare. Likewise, the SHAP (SHapley Additive exPlanations) framework developed by Lundberg and Lee [54] provides a unified, model-agnostic method for explaining individual predictions and has been widely adopted in clinical machine learning interpretability research.

Figure 12 titled “Top 15 Feature Importance (Random Forest)” illustrates the relative importance of features used by the Random Forest model in classifying physiotherapy treatment sequences for patients with bilateral idiopathic Carpal Tunnel Syndrome (CTS). Feature importance here was calculated using the mean decrease in impurity (MDI), measuring how much each feature contributed to reducing classification error across the decision trees in the ensemble.

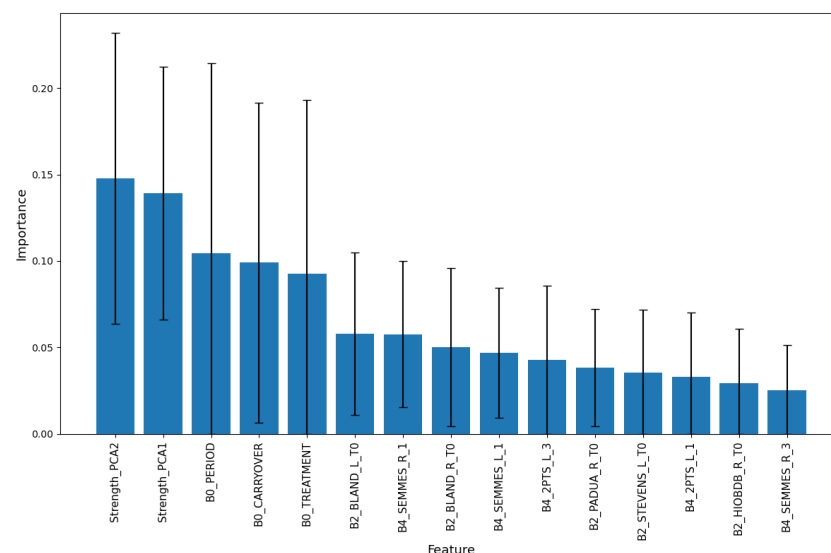


Figure 12. Top 15 feature importance scores identified by the Random Forest model, illustrating the relative contribution of key predictors to the classification of treatment sequences.

The two most influential features were the PCA-derived grip and pinch strength components: Strength_PCA2 and Strength_PCA1, contributing approximately 14.8% and 13.9% respectively to the model's decision-making. This highlights the predictive strength of biomechanical variables after dimensionality reduction. Other highly ranked features included B0_PERIOD, B0_CARRYOVER, and B0_TREATMENT, all related to treatment

timing, crossover effects, and intervention types, which are key factors in the study's clinical crossover design. Additional important contributors were diagnostic measures like B2_BLAND_L_T0 and B2_BLAND_R_T0 (Bland scale ratings at baseline) and sensory function tests such as B4_SEMMES_R_1, B4_SEMMES_L_1, and B4_2PTS_L_3, which assess fine touch and two-point discrimination abilities, closely linked to nerve impairment and recovery in CTS therapy.

Overall, the chart demonstrates the Random Forest model's ability to combine dimensionality-reduced strength indicators, clinical diagnostic scores, and treatment timing features into an interpretable framework for physiotherapy sequence prediction in clinical contexts.

Figure 13, titled “Top 15 Feature Importance (CatBoost),” displays the relative feature importance as determined by the CatBoost classifier for predicting appropriate physiotherapy sequences in bilateral CTS patients. In CatBoost, feature importance is determined based on the frequency and influence of feature splits during gradient boosting on decision trees.

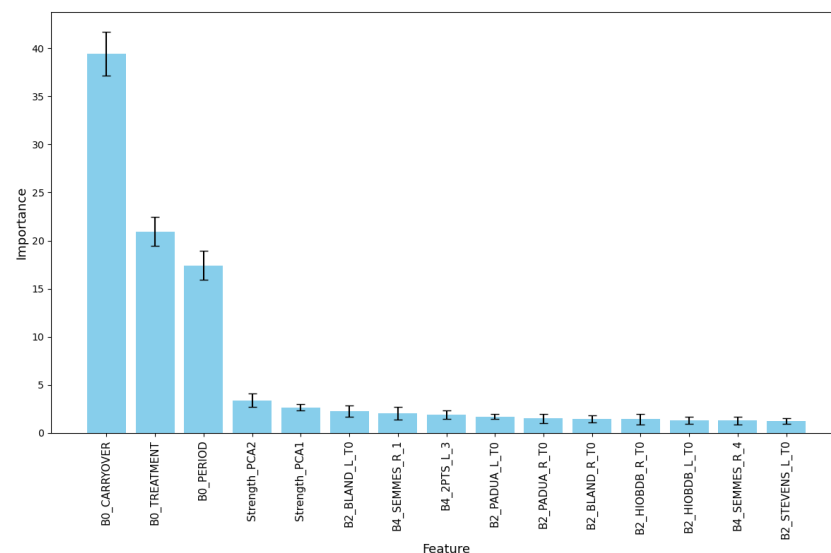


Figure 13. Top 15 feature importance scores identified by the CatBoost model, illustrating the relative contribution of key predictors to treatment sequence classification.

The most dominant feature, according to the chart, is B0_CARRYOVER, which alone contributes nearly 40% to decision-making. This feature indicates whether the treatment effects carried over from previous phases, a critical factor in crossover clinical studies. This highlights the significant role of carryover effects, often underestimated in physiotherapy studies. Following B0_CARRYOVER, the next influential features are B0_TREATMENT and B0_PERIOD, accounting for approximately 23% and 17% respectively, highlighting the importance of intervention type and timing.

Unlike Random Forest, PCA-derived strength features such as Strength_PCA2 and Strength_PCA1 ranked lower in CatBoost, suggesting that study-design variables contributed more to its decisions.

4.5. Ablation Analysis of Crossover-Related Variables

As part of the model evaluation and interpretability analysis, additional visualizations were incorporated to strengthen the assessment of CatBoost performance. First, the ROC-AUC curve was updated to include error bands representing ± 1 standard deviation across cross-validation folds, providing a clearer representation of the variability and stability of the classifier's performance. In addition, a feature importance analysis with error bars

was generated to illustrate the relative contribution of key predictors and the variability in their importance across folds, enabling a more robust interpretation of the model's learned feature relationships.

Furthermore, an ablation experiment was conducted in which the `B0_CARRYOVER` variable was removed from the CatBoost feature set. This step was performed to evaluate the extent to which the model's predictive performance was influenced by crossover-related information. After excluding this feature, the CatBoost model was retrained and re-evaluated, allowing a direct comparison with the original model. The resulting performance metrics showed a moderate reduction in predictive performance relative to the full-feature model, indicating that `B0_CARRYOVER` contributed to the model's accuracy but was not solely responsible for its predictive capability. Importantly, the classifier maintained strong discriminative ability, confirming that meaningful predictive signals are still captured from the remaining physiological and baseline features.

5. Security and Robustness

5.1. Adversarial Threat Model

In clinical machine learning applications, ensuring model robustness against adversarial attacks is critical for maintaining patient safety and system reliability. The adversarial threat model employed in this study was designed to simulate plausible perturbations in a clinical setting. Specifically, the attacks were constrained to minor modifications in the PCA-transformed feature space (`Strength_PCA1` and `Strength_PCA2`), representing subtle physiological variations that could realistically occur during data acquisition, such as sensor noise or recording inconsistencies.

The perturbation was crafted using a non-gradient-based method that gradually alters feature values toward the direction of the closest opposite-class sample. The attack operated under two strict constraints: a maximum Euclidean perturbation of 0.5 units and a step size of 0.01, ensuring that changes remained biologically plausible and imperceptible. This clinically grounded threat model focuses on evaluating whether machine learning models can maintain classification accuracy under minor but potentially misleading alterations in clinical input data.

5.2. CatBoost Adversarial Performance

In contrast to Random Forest, the CatBoost model exhibited slightly greater sensitivity to adversarial attack. The adversarial accuracy achieved by CatBoost was 94.01%, and the respective attack success rate was 5.99%. While overall accuracy for CatBoost was high, the increased attack success rate relative to Random Forest indicates a somewhat greater vulnerability to subtle input variations.

CatBoost, being a gradient-boosting tree model that is trained for categorical feature management, shows impressive baseline performance but may be more susceptible to attack from carefully tuned feature space directions. The gradient-based ensemble construction may introduce more pronounced decision boundaries, which are prone to more carefully crafted adversarial perturbations. Nevertheless, the robustness of CatBoost is still solid within the limited perturbation regime, and its intrinsic missing data and category split management offer important pragmatic benefits for clinical AI implementation.

5.3. Random Forest Adversarial Performance

The Random Forest classifier showed high robustness under the applied adversarial conditions. Following the perturbation process, the adversarial accuracy of the Random Forest classifier was 96.83%, which indicates only a minor reduction from the initial perfor-

mance. The attack success measure, which is the ratio of samples whose classification was modified after the attack, was fairly low at 3.17%.

This result indicates that Random Forest, being an ensemble-based approach that utilizes bagging and feature randomness, tends to counter small-scale adversarial shifts to a large extent. Most attacks were not capable of inducing misclassifications, indicating that the decision boundaries are fairly smooth and robust within the PCA feature space. However, a small percentage of boundary-proximity samples were successfully attacked, which highlights the need to further augment robustness by utilizing adversarial training or certification techniques in future research.

5.4. Comparative Robustness Analysis

The comparative analysis indicates that both Random Forest and CatBoost models demonstrate excellent adversarial robustness to low-magnitude, realistic clinical perturbations. Nevertheless, Random Forest performs better compared to CatBoost in rejecting adversarial misclassification for a given threat framework. Random Forest's inherent randomness, together with the ensemble averaging effect, can account for the robustness gap, while CatBoost's sequential optimization may lead to steeper decision regions vulnerable to small perturbations.

In all, these findings affirm that there is a need to incorporate robustness testing within clinical decision models' deployment. Adversarial training, ensemble smoothing, and certified robust processes are areas to be considered for future research to further advance the security stance of machine learning use within healthcare. Table 3 summarizes adversarial robustness results under clinically constrained PCA-space perturbations, reporting adversarial accuracy and attack success rates for Random Forest and CatBoost. Figure 14 provides a conceptual comparison between the adversarial attack process and the model-defense objective, illustrating how small latent-space perturbations can induce misclassification and how robustness evaluation mitigates this risk.

Table 3. Comparison of model robustness under adversarial perturbations, reporting adversarial accuracy and attack success rate.

Model	Adversarial Accuracy	Attack Success Rate	Interpretation
Random Forest	96.83%	3.17%	Highly robust; minimal attack success; smooth decision boundaries
CatBoost	94.01%	5.99%	Strong performance; moderately higher sensitivity to latent perturbations

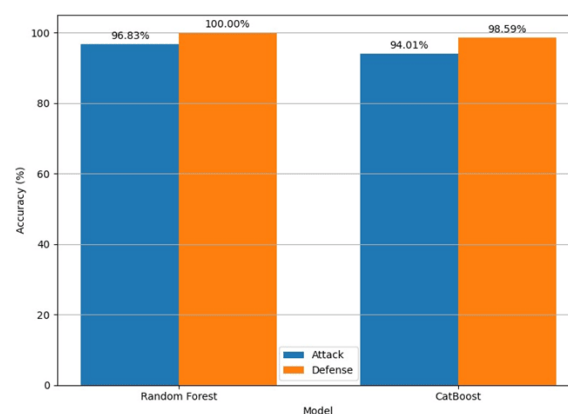


Figure 14. Illustration of the adversarial attack and defense process, demonstrating how minor feature perturbations can affect model predictions and how robustness testing helps maintain reliable classification.

6. Conclusions and Future Work

This study demonstrated that Random Forest and CatBoost models exhibit strong robustness against clinically realistic adversarial perturbations applied in the PCA-transformed feature space, maintaining an adversarial accuracy of 96.83% with a limited attack success rate of 3.17%. The results indicated that low-magnitude, latent-space perturbations targeting Strength_PCA1 and Strength_PCA2 have minimal impact on predictive performance, while explainability analyses confirmed that these features encode structured patterns relevant to bilateral carpal tunnel syndrome (CTS). Evaluation metrics, including F1-score, AUC, and precision–recall balance, remained consistently high before and after attacks, underscoring the models' stability. The study's contributions include the design of a biologically plausible adversarial threat model, a unified pipeline integrating PCA, ensemble learning, and controlled attacks, and a robustness evaluation framework extensible to other clinical domains.

6.1. Limitations

Several limitations of this study must be acknowledged to appropriately contextualize the findings and guide future research.

Randomized label as classification target. The most fundamental limitation is that the classification target, the treatment sequence label (S/M vs. M/S), was assigned by randomization in the source trial rather than derived from a clinical outcome measure. As a result, the models do not predict which treatment sequence is more effective for a given patient; they identify baseline feature patterns associated with sequence group membership within this specific dataset. The high classification accuracy observed may reflect structured feature-group associations present in this cohort rather than a generalizable clinical signal. This study is therefore exploratory and proof-of-concept in nature, and findings should not be interpreted as clinical recommendations.

Study design variables in feature importance. The CatBoost model assigned disproportionately high importance to B0_CARRYOVER (approximately 40%), a variable that reflects crossover study design rather than an intrinsic patient characteristic. In a prospective deployment scenario, such carryover information would not be available at the point of initial treatment assignment. This limits the direct clinical transferability of the CatBoost feature importance findings. Future work should conduct ablation experiments that exclude study-design variables to isolate the contribution of purely physiological predictors.

1. **Small sample size.** The final analytical dataset comprised 73 participants, further reduced after preprocessing, yielding a test set of 57 instances. While 10-fold cross-validation was employed to maximize the use of available data, the small sample size limits statistical power and increases the risk of performance estimates being sensitive to the specific train-test split. Replication on larger, independent cohorts is necessary to confirm the stability of the observed results.
2. **Restricted cohort generalizability.** The study cohort consisted exclusively of female, right-hand-dominant patients from a single geographic population in Brazil. While this homogeneity reduces confounding within the study, it substantially limits the generalizability of the findings to other patient populations, including male patients, left-hand-dominant individuals, and patients from different clinical or cultural contexts.
3. **Absence of outcome-based prediction.** The dataset contains multiple clinically meaningful outcome measures, including pain scores, BCTQ functional status, grip and pinch strength, and SF-12 quality-of-life scores, that were not used as prediction targets in this study. Future work should redesign the ML analysis around these outcome variables to produce findings with direct clinical relevance, such as predicting

which patients are most likely to benefit from one sequence over the other based on measurable functional improvement.

4. **Perturbation scope.** Adversarial robustness evaluation was confined to perturbations in the two PCA-derived strength features. This represents a narrow threat surface relative to the full feature space. Future robustness analyses should extend to gradient-based attacks (e.g., FGSM, PGD), multi-feature perturbations, and domain-shift scenarios involving corrupted or out-of-distribution clinical inputs.
5. **Model architecture.** Only tree-based ensemble classifiers were evaluated. Deep learning architectures, which may capture more complex nonlinear relationships in biomedical data, were not explored due to the small dataset size. Future studies with larger datasets should investigate neural network-based approaches alongside ensemble methods.

6.2. Future Directions

Future work should prioritize three directions. First, the ML framework should be redesigned around clinically derived outcome labels, such as change in pain score, BCTQ improvement, or grip strength recovery, to produce findings with direct clinical utility. Second, although an ablation experiment excluding B0_CARRYOVER was conducted in this study, future analyses should further investigate the contribution of additional study-design variables (e.g., B0_PERIOD) in order to more clearly isolate the predictive role of physiological features. Such analyses would help determine the extent to which classification performance is influenced by structural aspects of the crossover study design rather than intrinsic patient-related characteristics. Third, external validation on independent, larger, and more diverse CTS cohorts is necessary to assess the generalizability of the proposed pipeline before any consideration of clinical translation.

Author Contributions: Conceptualization, P.P.K., M.I.I., S.A.M., G.G. and F.A.; methodology, P.P.K., G.G. and M.I.I.; investigation, P.P.K., M.I.I., S.A.M., G.G. and F.A.; writing—original draft preparation, P.P.K., M.I.I. and G.G.; writing—review and editing, S.A.M., G.G. and F.A.; supervision, S.A.M., G.G. and F.A.; resources, P.P.K., M.I.I., S.A.M., G.G. and F.A.; data curation, P.P.K., M.I.I., S.A.M., G.G. and F.A.; visualization, P.P.K. and M.I.I. All authors have read and agreed to the published version of the manuscript.

Funding: The authors would like to express their heartfelt gratitude for the financial and technical support provided under the Assured and Trusted Digital Microelectronics Ecosystem (ADMETE) grant, BAA-FA8650-18-S-1201, funded by the Air Force Research Laboratory (AFRL). Partial support for this work was also received through funding to Wright State University, Dayton, Ohio, USA by the National Security Agency (NSA). This project adhered to CAGE Number: 4B991 and DUNS Number: 047814256.

Institutional Review Board Statement: Ethical review and approval were waived for this study because it is a secondary analysis of publicly available, de-identified data. The original study was conducted in accordance with the Declaration of Helsinki and approved by the Irmandade da Santa Casa de Londrina Ethics Committee (approval number: 3.276.439), as reported in the original dataset publication.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data used in this study are publicly available from the Mendeley Data repository. The dataset is titled “Dataset on Bilateral Idiopathic Carpal Tunnel Syndrome: Crossover Study of Two Combined Physiotherapeutic Treatment Methods on Chirurgical and Clinical Patients [4]”, and can be accessed at: <https://data.mendeley.com/datasets/47vj3yyb5v/1> (accessed on 29 January 2026). Additional data generated during the analysis are available from the corresponding author upon reasonable request.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Atroshi, I.; Gummesson, C.; Johnsson, R.; Ornstein, E.; Ranstam, J.; Rosen, I. Prevalence of carpal tunnel syndrome in a general population. *JAMA* **1999**, *282*, 153–158. [\[CrossRef\]](#)
2. Erickson, M.; Lawrence, M.; Jansen, C.W.S.; Coker, D.; Amadio, P.; Cleary, C.; Altman, R.; Beattie, P.; Boeglin, E.; Dewitt, J.; et al. Hand pain and sensory deficits: Carpal tunnel syndrome: Clinical practice guidelines linked to the international classification of functioning, disability and health from the academy of hand and upper extremity physical therapy and the academy of orthopaedic physical therapy of the American physical therapy association. *J. Orthop. Sport. Phys. Ther.* **2019**, *49*, CPG1–CPG85.
3. Dec, P.; Zyluk, A. Bilateral carpal tunnel syndrome—A review. *Neurol. I Neurochir. Pol.* **2018**, *52*, 79–83. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Georgeto, S.M.; Rogers, P.; Fernandes, K.B.P.; Andraus, R.A.C. Dataset on bilateral idiopathic carpal tunnel syndrome: Crossover study of two combined physiotherapeutic treatment methods on surgical and clinical patients. *Data Brief* **2024**, *56*, 110838. [\[CrossRef\]](#) [\[PubMed\]](#)
5. Dekel, S.; Papaioannou, T.; Rushworth, G.; Coates, R. Idiopathic carpal tunnel syndrome caused by carpal stenosis. *Br. Med. J.* **1980**, *280*, 1297–1299. [\[CrossRef\]](#) [\[PubMed\]](#)
6. O'Connor, D.; Marshall, S.C.; Massy-Westropp, N.; Pitt, V. Non-surgical treatment (other than steroid injection) for carpal tunnel syndrome. *Cochrane Database Syst. Rev.* **2003**, *2003*, CD003219. [\[CrossRef\]](#)
7. Hammer, W.I. The effect of mechanical load on degenerated soft tissue. *J. Bodyw. Mov. Ther.* **2008**, *12*, 246–256. [\[CrossRef\]](#)
8. AbdelRaouf, H.; Hasnaine, Q.R.; Fouda, M.M.; Fadlullah, Z.M.; Ibrahim, M.I. Paradigm Shift Toward Distributed Learning in IoT Intelligence: A Comprehensive Survey of Opportunities and Challenges. *IEEE Internet Things J.* **2026**, *1*. [\[CrossRef\]](#)
9. Fouda, M.M.; Fadlullah, Z.M.; Ibrahim, M.I.; Kato, N. Privacy-Preserving Data-Driven Learning Models for Emerging Communication Networks: A Comprehensive Survey. *IEEE Commun. Surv. Tutor.* **2025**, *27*, 2505–2542. [\[CrossRef\]](#)
10. AbdelRaouf, H.; Abouyoussef, M.; Fouda, M.M.; Ibrahim, M.I. Empowering AI-Driven Healthcare With Secure, Decentralized, and Privacy-Enhancing Adaptive Intelligence. *IEEE Internet Things J.* **2025**, *12*, 47223–47240. [\[CrossRef\]](#)
11. AbdelRaouf, H.; Abouyoussef, M.; Fouda, M.M.; Md Fadlullah, Z.; Ibrahim, M.I. Towards Decentralized, Secure, and Efficient Adaptive Learning for Robust Healthcare Monitoring. In *Proceedings of the ICC 2025—IEEE International Conference on Communications, Montreal, QC, Canada, 8–12 June 2025*; IEEE: Piscataway, NJ, USA, 2025; pp. 5469–5474. [\[CrossRef\]](#)
12. Esteva, A.; Robicquet, A.; Ramsundar, B.; Kuleshov, V.; DePristo, M.; Chou, K.; Cui, C.; Corrado, G.; Thrun, S.; Dean, J. A guide to deep learning in healthcare. *Nat. Med.* **2019**, *25*, 24–29. [\[CrossRef\]](#)
13. Paget, J. *Lectures on Surgical Pathology*; Longman: London, UK, 1853; Volume 1.
14. Ibrahim, I.; Khan, W.; Goddard, N.; Smitham, P. Carpal tunnel syndrome: A review of the recent literature. *Open Orthop. J.* **2012**, *6*, 69. [\[CrossRef\]](#)
15. Alfonso, C.; Jann, S.; Massa, R.; Torreggiani, A. Diagnosis, treatment and follow-up of the carpal tunnel syndrome: A review. *Neurol. Sci.* **2010**, *31*, 243–252. [\[CrossRef\]](#)
16. Abdelraouf, M.M.; Ahmed, A.I.; Elghitany, N.A. Sono-Elastography in evaluation of the median nerve in carpal tunnel syndrome patients. *QJM Int. J. Med.* **2021**, *114*, hcab106-052. [\[CrossRef\]](#)
17. Genova, A.; Dix, O.; Saefan, A.; Thakur, M.; Hassan, A.; Arguello, M.T. Carpal tunnel syndrome: A review of literature. *Cureus* **2020**, *12*, e7333. [\[CrossRef\]](#)
18. Joshi, A.; Patel, K.; Mohamed, A.; Oak, S.; Zhang, M.H.; Hsiung, H.; Zhang, A.; Patel, U.K. Carpal tunnel syndrome: Pathophysiology and comprehensive guidelines for clinical evaluation and treatment. *Cureus* **2022**, *14*, e27053. [\[CrossRef\]](#) [\[PubMed\]](#)
19. Lakhlani, D.; Shahoumian, T.A.; Curtin, C. Carpal tunnel syndrome diagnosis as a risk factor for falls. *Int. Orthop.* **2025**, *49*, 805–812. [\[CrossRef\]](#) [\[PubMed\]](#)
20. De Krom, M.; Kester, A.; Knipschild, P.; Spaans, F. Risk factors for carpal tunnel syndrome. *Am. J. Epidemiol.* **1990**, *132*, 1102–1110. [\[CrossRef\]](#)
21. Lakshminarayanan, K.; Shah, R.; Li, Z.M. Sex-related differences in carpal arch morphology. *PLoS ONE* **2019**, *14*, e0217425. [\[CrossRef\]](#) [\[PubMed\]](#)
22. Georgeto, S.M.; Andraus, R.A.C.; de Oliveira Júnior, E.; da Silva, R.A.; Ngomo, S.; Fernandes, K.B.P. Bilateral Idiopathic Carpal Tunnel Syndrome: Clinical-Functional Characterization and Efficacy of Two Combined Postoperative Physiotherapeutic Treatments. *Orthop. Surg.* **2023**, *15*, 1654–1663. [\[CrossRef\]](#)
23. Palmbergen, W.A.; Beekman, R.; Heeren, A.M.; van Nuenen, B.F.; Alleman, T.W.; Verstraete, E.; Jellema, K.; Verhagen, W.I.; Visser, L.H.; de Ruiter, G.C.; et al. Surgery versus corticosteroid injection for carpal tunnel syndrome (DISTRICTS): An open-label, multicentre, randomised controlled trial. *Lancet* **2025**, *405*, 2153–2163. [\[CrossRef\]](#)
24. Georgeto, S.; Picharski, G.; Andraus, R.; da Silva, R.; Ngomo, S.; Fernandes, K. Outcomes of bilateral carpal tunnel syndrome treatment—a systematic review and meta-analysis. *J. Plast. Reconstr. Aesthet. Surg.* **2022**, *75*, 3250–3259. [\[CrossRef\]](#)

25. Moser, F.; Muller, S.; Lie, T.; Langø, T.; Hoff, M. Automated segmentation of the median nerve in patients with carpal tunnel syndrome. *Sci. Rep.* **2024**, *14*, 16757. [\[CrossRef\]](#)
26. Coppieters, M.W.; Bartholomeeusen, K.E.; Stappaerts, K.H. Incorporating nerve-gliding techniques in the conservative treatment of cubital tunnel syndrome. *J. Manip. Physiol. Ther.* **2004**, *27*, 560–568. [\[CrossRef\]](#)
27. Coppieters, M.W.; Alshami, A.M. Longitudinal excursion and strain in the median nerve during novel nerve gliding exercises for carpal tunnel syndrome. *J. Orthop. Res.* **2007**, *25*, 972–980. [\[CrossRef\]](#)
28. Ho, C.Y.C.; Sole, G.; Munn, J. The effectiveness of manual therapy in the management of musculoskeletal disorders of the shoulder: A systematic review. *Man. Ther.* **2009**, *14*, 463–474. [\[CrossRef\]](#) [\[PubMed\]](#)
29. Hilton, C.B.; Milinovich, A.; Felix, C.; Vakharia, N.; Crone, T.; Donovan, C.; Proctor, A.; Nazha, A. Personalized predictions of patient outcomes during and after hospitalization using artificial intelligence. *NPJ Digit. Med.* **2020**, *3*, 51. [\[CrossRef\]](#) [\[PubMed\]](#)
30. Quattrocchi, S.; Russo, E.F.; Gatta, M.T.; Filoni, S.; Pellegrino, R.; Cangelmi, L.; Cardone, D.; Merla, A.; Perpetuini, D. Integrating Machine Learning with Robotic Rehabilitation May Support Prediction of Recovery of the Upper Limb Motor Function in Stroke Survivors. *Brain Sci.* **2024**, *14*, 759. [\[CrossRef\]](#)
31. Zhang, W. Enhancing Rehabilitation Assessment with Artificial Intelligence: A Comprehensive Investigation of Posture Quality Prediction Using Machine Learning. *ITM Web Conf.* **2025**, *70*, 02025. [\[CrossRef\]](#)
32. Rezaee, M.; Roshandel, H.; Rahimibarghani, S.; Rihani, T.S.S.; Mohammadyahya, E. Predictors of pain intensity in carpal tunnel syndrome: Development and validation of a model. *Clin. Neurol. Neurosurg.* **2024**, *243*, 108395. [\[CrossRef\]](#) [\[PubMed\]](#)
33. Franco-Moreno, A.; Madroñal-Cerezo, E.; de Ancos-Aracil, C.L.; Farfán-Sedano, A.I.; Muñoz-Rivas, N.; Bascuñana Morejón-Girón, J.; Ruiz-Giardin, J.M.; Álvarez-Rodríguez, F.; Prada-Alonso, J.; Gala-García, Y.; et al. Development of a Predictive Model of Occult Cancer After a Venous Thromboembolism Event Using Machine Learning: The CLOVER Study. *Medicina* **2024**, *61*, 18. [\[CrossRef\]](#)
34. Johnson, K.B.; Wei, W.Q.; Weeraratne, D.; Frisse, M.E.; Misulis, K.; Rhee, K.; Zhao, J.; Snowdon, J.L. Precision medicine, AI, and the future of personalized health care. *Clin. Transl. Sci.* **2021**, *14*, 86–93. [\[CrossRef\]](#)
35. Adeniran, I.A.; Efunniyi, C.P.; Osundare, O.S.; Abbulimen, A.O. Data-driven decision-making in healthcare: Improving patient outcomes through predictive modeling. *Int. J. Sch. Res. Multidiscip. Stud.* **2024**, *5*, 59–67. [\[CrossRef\]](#)
36. Shi, X.; Yu, T.; Yuan, Y.; Wang, D.; Cui, J.; Bai, L.; Zheng, F.; Dai, X.; Du, R.; Chen, Z.; et al. Multimodal deep learning for grading carpal tunnel syndrome: A multicenter study in China. *Acad. Radiol.* **2025**, *32*, 4705–4723. [\[CrossRef\]](#)
37. Tsamis, K.I.; Kontogiannis, P.; Gourgiotis, I.; Ntabos, S.; Sarmas, I.; Manis, G. Automatic electrodiagnosis of carpal tunnel syndrome using machine learning. *Bioengineering* **2021**, *8*, 181. [\[CrossRef\]](#) [\[PubMed\]](#)
38. Hoogendam, L.; Bakx, J.A.C.; Souer, J.S.; Slijper, H.P.; Andrinopoulou, E.R.; Selles, R.W.; Hand Wrist Study Group. Predicting clinically relevant patient-reported symptom improvement after carpal tunnel release: A machine learning approach. *Neurosurgery* **2022**, *90*, 106–113. [\[CrossRef\]](#)
39. Park, D.; Kim, B.H.; Lee, S.E.; Kim, D.Y.; Kim, M.; Kwon, H.D.; Kim, M.C.; Kim, A.R.; Kim, H.S.; Lee, J.W. Machine learning-based approach for disease severity classification of carpal tunnel syndrome. *Sci. Rep.* **2021**, *11*, 17464. [\[CrossRef\]](#)
40. Harrison, C.J.; Geoghegan, L.; Sidey-Gibbons, C.J.; Stirling, P.H.C.; McEachan, J.E.; Rodrigues, J.N. Developing machine learning algorithms to support patient-centered, value-based carpal tunnel decompression surgery. *Plast. Reconstr. Surg. Glob. Open* **2022**, *10*, e4279. [\[CrossRef\]](#)
41. Elseddik, M.; Mostafa, R.R.; Elashry, A.; El-Rashidy, N.; El-Sappagh, S.; Elgamal, S.; Aboelfetouh, A.; El-Bakry, H. Predicting CTS diagnosis and prognosis based on machine learning techniques. *Diagnostics* **2023**, *13*, 492. [\[CrossRef\]](#) [\[PubMed\]](#)
42. Elseddik, M.; Alnowaiser, K.; Mostafa, R.R.; Elashry, A.; El-Rashidy, N.; Elgamal, S.; Aboelfetouh, A.; El-Bakry, H. Deep learning-based approaches for enhanced diagnosis and comprehensive understanding of carpal tunnel syndrome. *Diagnostics* **2023**, *13*, 3211. [\[CrossRef\]](#) [\[PubMed\]](#)
43. Sorin, V.; Soffer, S.; Glicksberg, B.S.; Barash, Y.; Konen, E.; Klang, E. Adversarial attacks in radiology—A systematic review. *Eur. J. Radiol.* **2023**, *167*, 111085. [\[CrossRef\]](#)
44. Finlayson, S.G.; Bowers, J.D.; Ito, J.; Zittrain, J.L.; Beam, A.L.; Kohane, I.S. Adversarial attacks on medical machine learning. *Science* **2019**, *363*, 1287–1289. [\[CrossRef\]](#)
45. Kaviani, S.; Han, K.J.; Sohn, I. Adversarial attacks and defenses on AI in medical imaging informatics: A survey. *Expert Syst. Appl.* **2022**, *198*, 116815. [\[CrossRef\]](#)
46. Tsai, M.J.; Lin, P.Y.; Lee, M.E. Adversarial Attacks on Medical Image Classification. *Cancers* **2023**, *15*, 4228. [\[CrossRef\]](#) [\[PubMed\]](#)
47. Bhatta, N.P.; Al Majmaie, S.; Amsaad, F. Feature Analysis and Model Evaluation for Classification of Hardware Trojans. In *Proceedings of the 2024 IEEE Physical Assurance and Inspection of Electronics (PAINE)*, Huntsville, AL, USA, 12–14 November 2024; IEEE: Piscataway, NJ, USA, 2024; pp. 1–6.
48. Jolliffe, I.T. *Principal Component Analysis*; Springer Series in Statistics; Springer: Berlin/Heidelberg, Germany, 2002. [\[CrossRef\]](#)
49. Abdi, H.; Williams, L.J. Principal component analysis. *Wiley Interdiscip. Rev. Comput. Stat.* **2010**, *2*, 433–459. [\[CrossRef\]](#)
50. Bergstra, J.; Bengio, Y. Random search for hyper-parameter optimization. *J. Mach. Learn. Res.* **2012**, *13*, 281–305.

51. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Müller, A.; Nothman, J.; Louppe, G.; et al. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
52. Kohavi, R. A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Proceedings of the 14th International Joint Conference on Artificial Intelligence (IJCAI), Montreal, QC, Canada, 20–25 August 1995*; Morgan Kaufmann Publishers Inc.: San Francisco, CA, USA, 1995; Volume 2, pp. 1137–1143.
53. Molnar, C. *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable*; Lean Publishing: Victoria, BC, Canada, 2020.
54. Lundberg, S.M.; Lee, S.I. A Unified Approach to Interpreting Model Predictions. In *Proceedings of the Advances in Neural Information Processing Systems, Long Beach, CA, USA, 4–9 December 2017*; Curran Associates Inc.: Red Hook, NY, USA, 2017; Volume 30.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.