applied sciences

MDPI

Article

Prediction of Electrophysiological Severity and Carpal Tunnel Syndrome Instrument Changes After Carpal Tunnel Release Using Machine Learning Model

Atsuyuki Inui [1,*], Fumiaki Takase [2], Stefano Lucchina [3,4,5] and Takako Kanatani [2]

[1] Department of Orthopaedic Surgery, Kobe University Graduate School of Medicine, Kobe 650-0017, Japan
[2] Department of Orthopaedic Surgery, Kobe Rosai Hospital, Kobe 651-0053, Japan
[3] Locarno Hand Center, 6600 Locarno, Switzerland
[4] Surgical Department, Hand Unit EOC, Locarno's Regional Hospital, 6600 Locarno, Switzerland
[5] Service of Orthopaedics and Traumatology, Department of Surgery, EOC, 6900 Lugano, Switzerland
* Correspondence: ainui@med.kobe-u.ac.jp; Tel.: +81-78-382-5985

Abstract: Introduction: The severity of carpal tunnel syndrome (CTS) is evaluated by electrophysiological examination as well as a patient-oriented questionnaire. We hypothesized that machine learning could predict postoperative electrophysiological severity as well as the scores of patient-oriented questionnaires. In this study, we developed machine learning models to predict postoperative changes in electrophysiological severity and changes in the Carpal Tunnel Syndrome Instrument (CTSI). Materials and Methods: Data from four hundred and twenty hands of individuals who had been diagnosed with CTS and undergone carpal tunnel release were used. The features used for the machine learning model were preoperative age, gender, distal motor latency (DML) value, sensory nerve conduction velocity (SCV) value, preoperative electrophysiological severity stage, CTSI-SS value, and CTSI-FS value. Logistic Regression (LR), ElesticNet (EN), Support Vector Machine (SVM), Random Forest (RF), and LightGBM (LGBM) were used as machine learning algorithms. A machine learning model was created to binary classify the electrophysiologic severity at one year postoperatively. In the second experiment, regression models were created to predict the change in CTSI-SS and CTSI-FS at one year postoperatively. Results: In the electrophysiological severity classification model, LGBM showed the highest score (AUC = 0.802). Preoperative DML, age, and preoperative electrophysiological severity were important factors for model prediction. RF model showed the best performance. In the regression model predicting the change in CTSI-SS or CTSI-FS (RMSE: 0.418, 0.333, respectively), preoperative age and CTSI-SS or CTSI-FS scores were important factors for model prediction. Conclusions: The machine learning model can predict postoperative electrophysiological severity and CTSI score with high accuracy.

Keywords: CTS; machine learning; nerve conduction study

check for updates

Academic Editor: Jing Jin

Received: 22 December 2024
Revised: 29 January 2025
Accepted: 3 February 2025
Published: 10 February 2025

Citation: Inui, A.; Takase, F.; Lucchina, S.; Kanatani, T. Prediction of Electrophysiological Severity and Carpal Tunnel Syndrome Instrument Changes After Carpal Tunnel Release Using Machine Learning Model. *Appl. Sci.* **2025**, *15*, 1812. https://doi.org/10.3390/app15041812

Copyright: © 2025 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) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

CTS is the most common entrapment neuropathy of the upper extremity and is more common in women, with a peak age range of 40–60 years [1–3]. Diagnoses are subjectively made based on clinical findings of sensory deficit and/or thenar atrophy [4], while they are made objectively via NCS and ultrasound examination [5]. NCS is useful not only in diagnosing CTS but also in evaluating nerve degeneration. However, NCS is an objective method of assessment and does not always correspond to clinical symptoms [6]. Moreover,

electrophysiological recovery does not always support improvements in clinical outcome. Therefore, a patient-orientated questionnaire that assesses patient-perceived symptoms and quality of life (QOL) deterioration is important in determining the best treatment strategy and evaluating the efficacy of the treatment. Performing an NCS and patient-orientated questionnaire pre- and postoperatively is ideal; however, it requires time, comes with high costs, and cooperation from patients; furthermore, in the case of NCSs, patients may experience discomfort. If the postoperative changes determined by the NCS and patient-orientated questionnaires are predicted by preoperative measurements, these problems would be overcome; furthermore, it would also suggest that these approaches are a practical choice for selecting the treatment method.

The role of nerve conduction studies (NCSs) in carpal tunnel syndrome (CTS) is not only to diagnose CTS but also to objectively evaluate the severity of nerve damage. We have previously reported the usefulness of preoperative evaluation and postoperative assessment using the electrophysiological severity classification (Stages 1–5) based on parameters of distal motor latency (DML) and sensory nerve conduction velocity (SCV) of the median nerve [7]. However, patients' main complaints are pain, numbness, and difficulty with pinching movements, which do not correlate with electrophysiological severity [2]. Therefore, postoperative assessments focusing on patient-oriented outcomes are also important, as postoperative electrophysiologic severity assessment alone is insufficient to determine treatment efficacy. The negative aspects of postoperative NCS are the costs and patient discomfort during the procedure. Also, the use of a patient-oriented questionnaire requires patient cooperation. Therefore, we consider if postoperative changes in NCS and patient-oriented questionnaires could be predicted from preoperative assessments, as this would be ideal in determining the treatment strategy. Currently, machine learning applications in the medical field are expanding [8,9]. Further improvements in the accuracy of diagnosis and the capability of the prediction of prognoses by machine learning may help us avoid human error in diagnosis and in determining the treatment method. We hypothesized that machine learning, which is a field of artificial intelligence (AI), could predict postoperative electrophysiological severity as well as the scores of patient-oriented questionnaires. In this study, we developed machine learning models to predict postoperative changes in electrophysiological severity, as well as the Carpal Tunnel Syndrome Instrument (CTSI) of the Japanese Society for Surgery of the Hand (JSSH) version; CTSI-JSSH [10].

2. Materials and Methods

2.1. Data Collection

Before the present investigation, we obtained approval from the Ethics Committees of our institution (Approved number: 4-04). Written informed consent was obtained from each patient after the oral explanations of the details of the present study were given. None of the eligible patients declined enrolment.

Four hundred and twenty hands of patients diagnosed with CTS at our hospital from March 2013 to September 2022 were included. The diagnosis of carpal tunnel syndrome was made by a hand surgeon based on numbness in the median nerve area, carpal tunnel Tinel's signs, and electrophysiology results. The severity of carpal tunnel syndrome was classified following previous reports [7,11]. The patients underwent open carpal tunnel release (CTR) with small skin incisions and underwent NCS and CTSI-JSSH scoring before and one year after surgery. The average age of the patients was 68.3 $\pm$ 12.1 years (26–97 years), with 102 male and 318 female hands. Cases with diabetic peripheral neuropathy or cervical spine disorders were excluded. For NCS, the distance between the stimulus and derivation electrodes was set to 7 cm, and the DML of the abductor pollicis brevis was recorded with maximum electrical stimulation at the wrist joint. SCV was measured by the retrograde

method with maximum electrical stimulation at the wrist joint and derivation at the index finger. The distance between the stimulus and derivation electrodes was 14 cm. The preoperative and one-year postoperative values were classified into Stages 1–5 using the electrophysiological severity classification with the cutoff values as DML $\geqq$ 4.5 ms and SCV < 40.0 m/s [7,11]. All tests were performed by the same examiner using MEB2200 (Nihon Kohden, Tokyo, Japan).

CTSI-JSSH is composed of the severity score of symptoms (CTSI-SS) and the scale of functional status (CTSI-FS). CTSI-SS is calculated as the average of the contents of 11 items evaluated from 1 to 5 points, with mild being 1 point and most severe being 5 points. CTSI-FS is used for finger function evaluation, and the average of 8 questions is calculated, with mild being 1 point and most severe being 5 points. The CTSI-JSSH questionnaire was completed by the patients themselves outside the examination room.

2.2. Machine Learning Model Training

For machine learning, we used the scikit-learn library (ver. 1.2.2) of Python. To create a machine learning dataset, undetectable electrophysiological data were preprocessed as follows. Cases where DML could not be detected were replaced with DML = 20 ms. Cases where SCV could not be detected were replaced with SCV = 0 m/s. In total, 70% of all data were used as training data and 30% were used as test data. The features used for the machine learning model are summarized in Table 1.

Table 1. The features used for the machine learning model.

Abbreviation	Meaning
Age	Age at the time of surgery
IsFemale	Gender: female or male
preDML	Pre-operative DML
preSCV	Pre-operative SCV
prestage	Pre-operative electrophysiological severity stage
preSS	Pre-operative CTSI-SS score
preFS	Pre-operative CTSI-FS score

The first machine learning model was created to binary classify electrophysiologic severity at one year postoperatively. Logistic Regression (LR), Support Vector Machine (SVM), Random Forest (RF), and LightGBM (LGBM) were used as machine learning algorithms. The characteristics of each machine learning model are as follows. Logistic Regression (LR): a linear model for binary classification, estimating probabilities using a logistic function. Support Vector Machine (SVM): a supervised learning model finding the optimal hyperplane to classify data points. Random Forest (RF): an ensemble learning method using multiple decision trees to improve classification and regression performance. LightGBM (LGBM): a gradient boosting framework that uses tree-based learning algorithms for fast, accurate results. In the second experiment, LR, Elastic Net, SVM, RF, and LGBM were used as regression models to predict the change in CTSI-SS and CTSI-FS at one year postoperatively.

2.3. Statistical Analysis

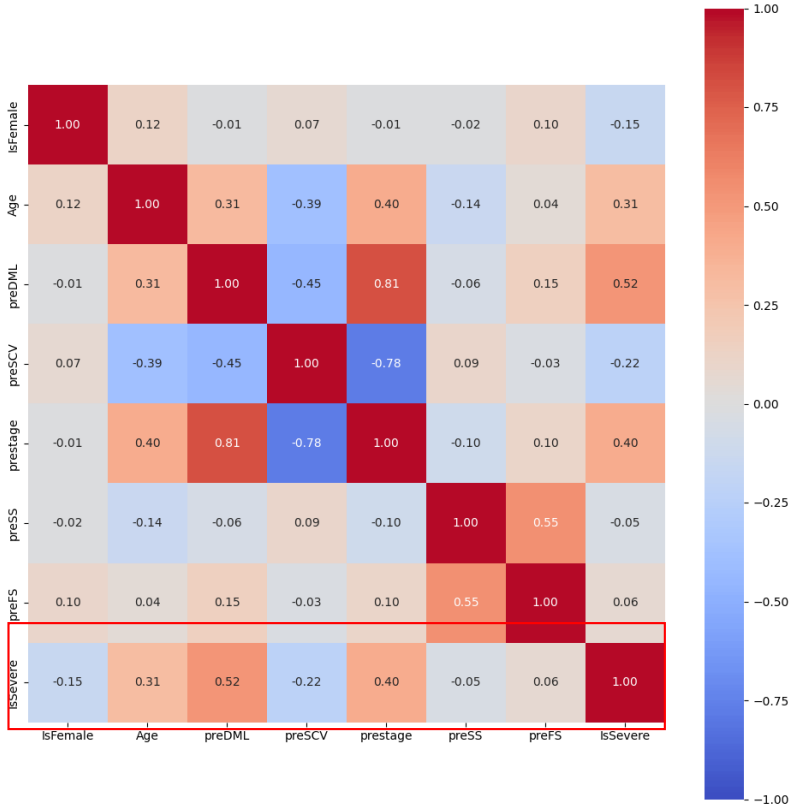
Correlation coefficients between each parameter are displayed in the heatmap. In this heatmap, the correlation coefficients are plotted in the range of −1.0 to 1.0, with warm colors showing a positive correlation and cold colors showing a negative correlation. Pearson's correlation coefficient test was performed and items with correlation coefficients greater than 0.4 or less than −0.4 were considered correlated. For statistical examination, a chi-square test and paired *t*-test were performed on the distribution of electrophysiological severity, CTSI-SS scores, and CTSI-FS scores. For electrophysiological severity, logistic regression analysis was performed after grouping Stages 4 and 5 into the severe group and Stages 1–3 into the mild and moderate group. Multivariate analysis was performed for statistically significant variables to determine changes in CTSI-SS and CTSI-FS. Statistical analysis was performed using the Statsmodels library (ver. 0.14.1) of Python (ver. 3.10.12), and $p < 0.05$ was set as statistically significant.

For machine learning classification model evaluation, the accuracy, recall (sensitivity), precision, and f-measure, as well as the specificity for each model, were calculated from the confusion matrix obtained from the test data. The characteristics of each evaluation method are as follows. Accuracy: the proportion of correctly predicted instances among all instances. Recall: the proportion of actual positives correctly identified by the model. Precision: the proportion of positive predictions that are truly positive. F-Measure: the harmonic mean of precision and recall. Specificity: the proportion of actual negatives correctly identified by the model. The AUC (Area Under the Curve) was calculated from the Receiver Operating Characteristic (ROC) curve. The regression models were evaluated using the Root Mean Square Error (RMSE), the Mean Absolute Error (MAE), and the coefficient of determination (R^2). Machine learning model results were presented in the form of a mean (95% confidence interval). As a method of explainable AI (XAI), the important features in the model were visualized using feature importance and SHAP (Shapley Additive Explanation) scores. Feature importance is a concept that shows how important each feature is in improving the accuracy of the model. The SHAP score is a score that quantifies how much each feature has contributed to the prediction result of the model based on game theory.

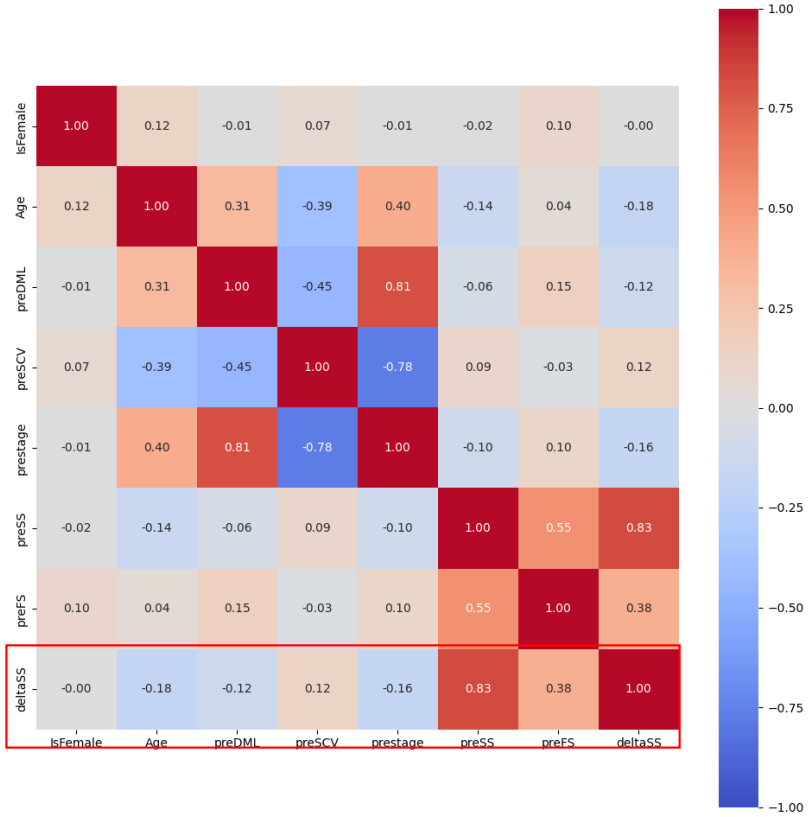
3. Results

3.1. Statistical Analysis Results

The correlation coefficients between each parameter are shown in Figure 1 as a heatmap, with the bottom column of the heatmap representing the correlation between the objective and explanatory variables. Postoperative electrophysiological severity was positively correlated with preoperative DML and preStage (correlation coefficient: *p*-value = 0.52: 3.85×10^{-30}, 0.40: 9.08×10^{-18}, respectively) (Figure 1a). The postoperative CTSI-JSSH-SS change correlated positively with the preoperative CTSI-JSSH-SS score (correlation coefficient: *p*-value = 0.83: 1.33×10^{-106}) (Figure 1b). The postoperative CTSI-JSSH-FS change correlated positively with the preoperative CTSI-JSSH-FS and CTSI-JSSH-SS score (correlation coefficient: *p*-value = 0.87: 2.47×10^{-133}, 0.48: 2.71×10^{-25}, respectively) (Figure 1c).

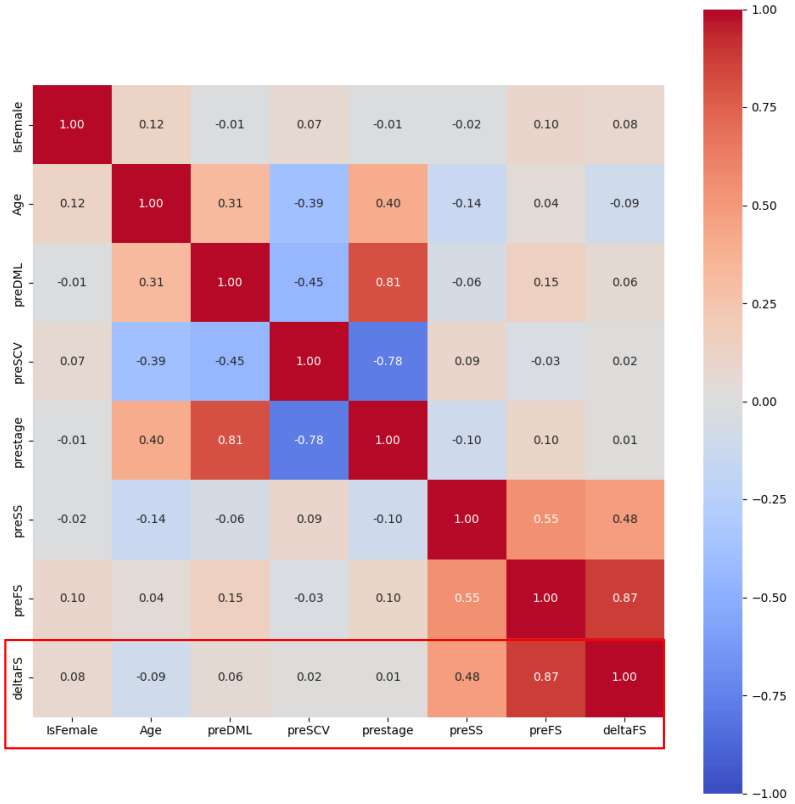


(**a**) Correlation between electrophysiological severity (IsSevere) and each parameter.



(**b**) Correlation between the change in CTS − SS (deltaSS) and each parameter.

Figure 1. *Cont.*



(**c**) Correlation between the change in CTSI − FS (deltaFS) and each parameter.

Figure 1. Heatmap of the correlation coefficients between each parameter. The bottom column of the heatmap represents the correlation between the objective and explanatory variables. Warm colors show a strong positive correlation and cold colors have a strong negative correlation.

According to the results of the x-square test using Table 2, the postoperative electrophysiological severity was significantly improved compared to the preoperative value. ($p = 1.0 \times 10^{-4}$). According to the results of a paired *t*-test pre- and postoperative CTSI scores, the mean CTSI-JSSH-SS was 2.52 ± 0.85 preoperatively and 1.33 ± 0.51 at one year postoperatively, indicating a statistically significant decrease. ($p = 2.54 \times 10^{-76}$). The CTSI-JSSH-FS score was 2.40 ± 0.96 preoperatively and 1.42 ± 0.63 at one year postoperatively, showing a statistically significant decrease ($p = 2.97 \times 10^{-60}$). In a logistic regression analysis comparing the mild and moderate, and severe electrophysiologic severity groups at one year postoperatively, gender, age, and preoperative DML were statistically significant variables ($p = 2.60 \times 10^{-5}$, $p = 2.99 \times 10^{-6}$, $p = 3.26 \times 10^{-11}$, respectively). In a multivariate analysis of changes in CTSI-JSSH-SS scores before and after surgery, preoperative CTSI-JSSH-FS and preoperative CTSI-JSSH-SS scores were statistically significant variables ($p = 1.31 \times 10^{-8}$, $p = 0.022$, respectively). In a multivariate analysis of pre-and postoperative CTSI-JSSH-FS changes, preoperative CTSI-JSSH-FS, age, and preoperative SCV were statistically significant variables ($p = 1.59 \times 10^{-76}$, $p = 6.65 \times 10^{-7}$, $p = 0.026$, respectively). The t-values for each item in the multiple regression analysis are summarized in Table 3 as an indication of the contribution of each item in predicting prognosis.

Table 2. Distribution of pre-and postoperative electrophysiological severity.

	Pre-Surgery (Hands)	Post-Surgery (Hands)
Stage 1	3	113
Stage 2	10	108
Stage 3	59	89
Stage 4	149	68
Stage 5	198	38
unclassified	1	4

Table 3. The t-values for each item in the multiple regression analysis are summarized as an indicator of the contribution of each item in predicting prognosis.

	Change In CTSI-JHHS-FS Score	Change In CTSI-JHHS-SS Score
Is Female	1.17	0.46
Age	−1.37	−4.87
preDML	0.11	0.22
preSCV	−0.91	−2.12
preStage	−1.20	−1.75
preSS	25.8	−1.35
preFS	−2.4	32.0

3.2. Results of Machine Learning Model

3.2.1. Classification Model

The performance of the LR, SVM, RF, and LGBM models in the task of classifying the electrophysiological severity of the test data is shown in Table 4. The AUC obtained from the ROC curve was 0.754 (0.752–0.757) for LR, 0.765 (0.763–0.768) for SVM, 0.764 (0.761–0.766) for RF, and 0.802 (0.799–0.804) for LGBM (Figure 2). The LGBM model showed the highest score in all evaluation metrics. When visualizing the feature importance of the LGBM model, preoperative DML, age, and preoperative electrophysiological severity stage were important factors. The SHAP score for the same LGBM model showed that preoperative DML, age, and gender were important factors (Figure 3).

Table 4. Model performance in electrophysiological severity classification task.

	LR	SVM	RF	LGBM
Accuracy	0.819 (0.817–0.821)	0.811 (0.809–0.812)	0.829 (0.827–0.831)	0.851 (0.849–0.853)
Recall (Sensitivity)	0.619 (0.614–0.623)	0.671 (0.666–0.675)	0.625 (0.621–0.639)	0.697 (0.692–0.701)
Precision	0.667 (0.663–0.672)	0.630 (0.626–0.635)	0.694 (0.690–0.699)	0.724 (0.720–0.728)
F-measure	0.640 (0.636–0.643)	0.647 (0.644–0.651)	0.656 (0.652–0.659)	0.708 (0.704–0.711)
Specificity	0.890 (0.888–0.892)	0.861 (0.859–0.863)	0.902 (0.900–0.904)	0.906 (0.904–0.907)
Best Parameters of Model	Model intercept = −1.602	C = 100 Gamma = 0.001	Max depth = 3 Numbers of estimators = 300	Objective = binary Metric = binary error Num_leaves = 6

The performance of the LR, SVM, RF, and LGBM models in the task of classifying electrophysiological severity. LR: Logistic Regression; SVM: Support Vector Machine; RF: Random Forest; LGBM: LightGBM.

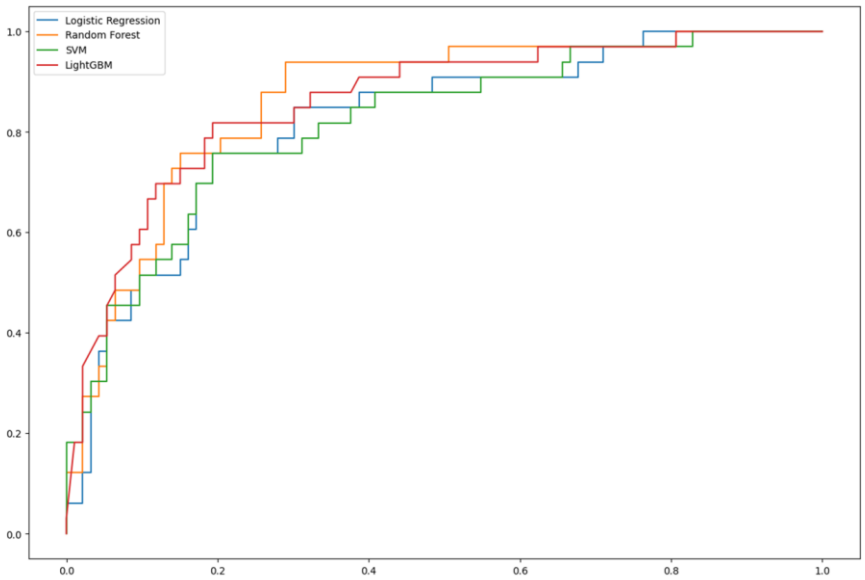
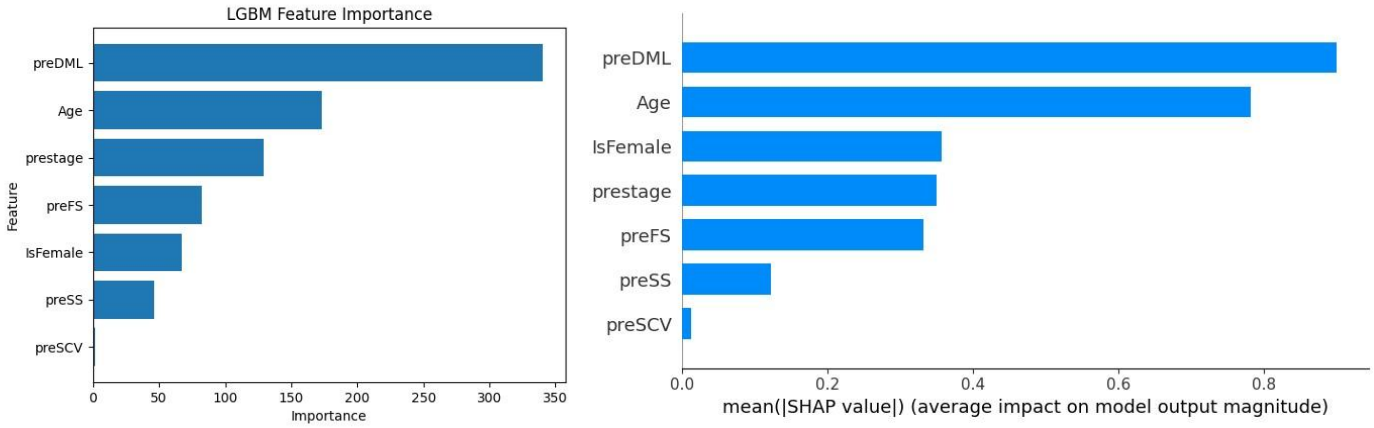


Figure 2. The AUC obtained from the ROC curve was 0.754 (0.752–0.757) for LR, 0.765 (0.763–0.768) for SVM, 0.764 (0.761–0.766) for RF, and 0.802 (0.799–0.804) for LGBM.



(a) (b)

Figure 3. (**a**) Feature importance; (**b**) SHAP scores of the LGBM model in predicting electrophysiological severity one year after the surgery.

3.2.2. Regression Model

The performance of the regression model in predicting the change in CTSI-JSSH-SS scores before and after surgery is shown in Table 5. The RF model showed the best RMSE and R^2 score among the five models. Visualizing the feature importance of the RF model, preoperative CTSI-JSSH-SS score, preoperative CTSI-JSSH-FS score, and age were important factors. Similarly, preoperative CTSI-JSSH-SS score, preoperative CTSI-JSSH-FS score, and preoperative DML were important factors for the SHAP score of the RF model (Figure 4). The performance of the model in predicting changes in CTSI-JSSH-FS before and after surgery is shown in Table 6. The RF model had the best RSME and R^2 scores of the five models. Visualization of feature importance and SHAP values showed that preoperative CTSI-JSSH-FS score, age, and preoperative CTSI-SS score were important factors (Figure 5).

Table 5. Model performance in the task of detecting the change in CTSI-SS score.

	LR	EN	SVM	RF	LGBM
RMSE	0.486 (0.478–0.494)	0.485 (0.476–0.494	0.487 (0.478–0.495)	0.418 (0.410–0.426)	0.475 (0.464–0.485)
MAE	0.337 (0.332–0.342)	0.336 (0.330–0.341)	0.315 (0.309–0.320)	0.285 (0.280–0.290)	0.267 (0.261–0.272)
R2	0.688 (0.679–0.697)	0.685 (0.675–0.695)	0.683 (0.672–0.694)	0.768 (0.760–0.776)	0.702 (0.691–0.713)
Best Parameters of Model	Model intercept = 1.39	Alpha = 0.004 L1_ratio = 0.00	C = 0.464 Gamma = 0.021	Max depth = 5 Numbers of estimators = 300	Learning rate −0.073 Max depth = 3

The performance of the LR, EN, SVM, RF, and LGBM models for the task of predicting the change in the CTSI-JSSH-SS score. LR: Logistic Regression; EN: Elastic Net; SVM: Support Vector Machine; RF: Random Forest; LGBM: LightGBM.

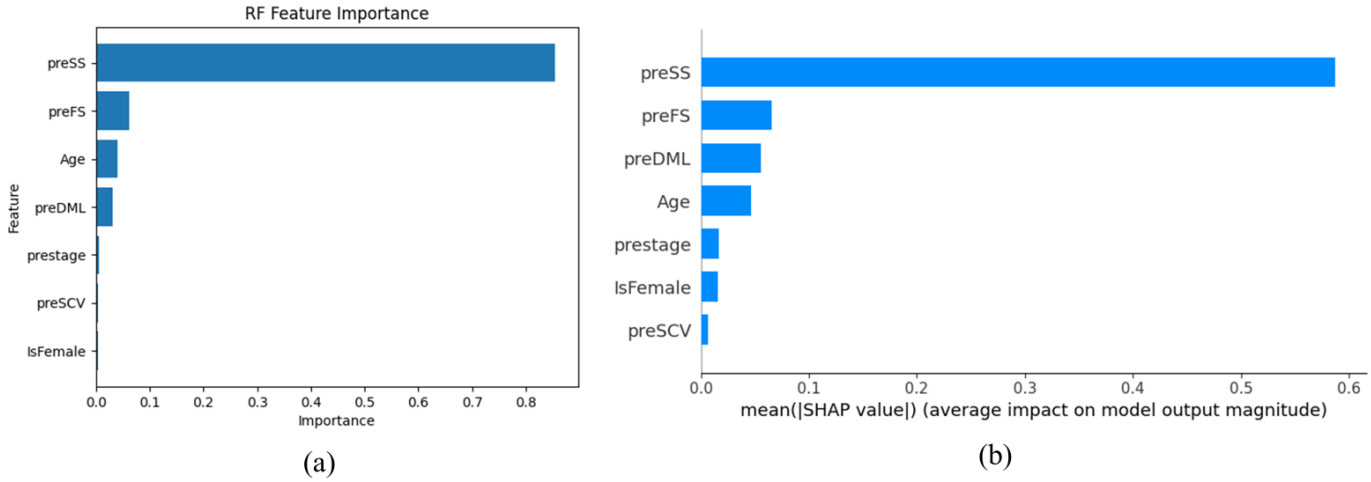
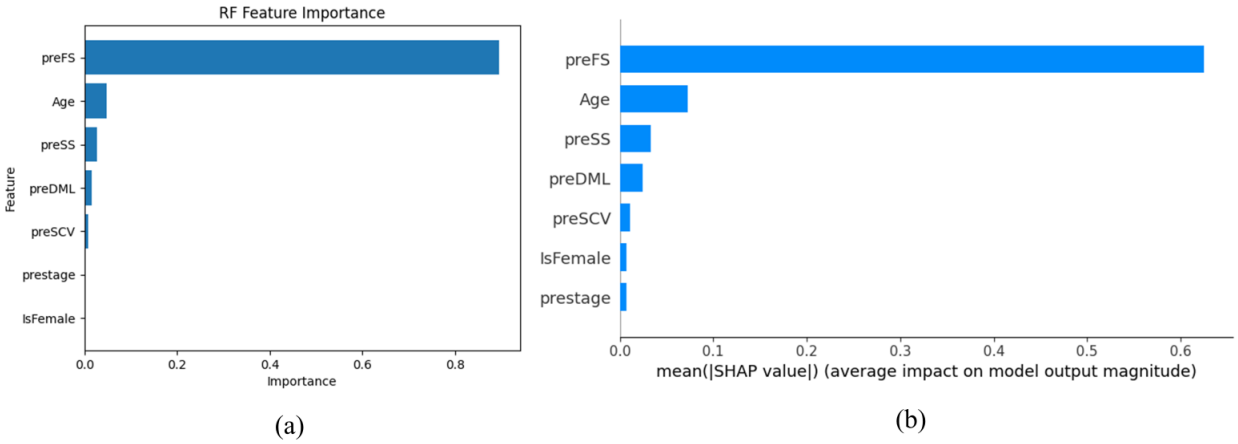


Figure 4. (**a**) Feature importance, (**b**) SHAP scores of the RF model in predicting changes in CTSI-JSSH-SS scores one year after the surgery.

Table 6. Model performance in the task of detecting the change in CTSI-FS scores.

	LR	EN	SVM	RF	LGBM
RMSE	0.405 (0.403–0.407)	0.404 (0.402–0.406)	0.399 (0.397–0.401)	0.333 (0.332–0.335)	0.375 (0.373–0.377)
MAE	0.289 (0.287–0.290)	0.288 (0.287–0.289)	0.271 (0.270–0.273)	0.238 (0.236–0.239)	0.234 (0.233–0.235)
R^2	0.771 (0.769–0.774)	0.773 (0.771–0.776)	0.779 (0.776–0.781)	0.846 (0.844–0.847)	0.803 (0.801–0.806)
Best Parameters of Model	Model intercept = 1.22	Alpha = 0.0021 learning_ ratio = 0.00	C = 0.464 Gamma = 0.022	Max depth = 5 Numbers of estimators = 100	Objective = Mean absolute error Learning rate −0.01 Max depth = 7

The performance of the LR, EN, SVM, RF, and LGBM models in the task of predicting the change in CTSI-JSSH-FS scores. LR: Logistic Regression; EN: Elastic Net; SVM: Support Vector Machine; RF: Random Forest; LGBM: LightGBM.



(a) (b)

Figure 5. (**a**) Feature importance; (**b**) SHAP scores of the RF model in predicting the change in CTSI-JSSH-FS scores one year after the surgery.

4. Discussion

This study established a machine learning model to predict postoperative changes in electrophysiological severity or CTSI scores after surgery for carpal tunnel syndrome. The LightGBM model was effective in predicting electrophysiological severity and the Random Forest model was effective in predicting postoperative changes in CTSI scores.

In recent years, there has been an increasing number of reports using AI models for the diagnosis and treatment of CTS [12–14]. Most of them report on the utility of AI models in the diagnosis of CTS. Park et al. [12] proposed an AI model to classify the electrophysiological severity of CTS as mild, moderate, and severe based on age, body size, questionnaire responses, and ultrasound findings. Among seven machine learning models, eXtremeGradient Boosting (XGB) was the most accurate [12]. Tsamis et al. [13] created a machine learning model by generating multiple parameters from DML, motor nerve conduction velocity, SCV, compound motor action potential, and sensory nerve action potential waveforms. They reported that the accuracy of this model in diagnosing CTS was higher than that of conventional NCS. Shinohara et al. [14] reported that deep learning on the ultrasound image of the median nerve provided a highly accurate diagnosis without nerve conduction studies. On the other hand, there are a few papers regarding the changes in NCS and patient-oriented questionnaires after CTS treatment. Elseddick et al. [15] used CTSI, NCS, and ultrasound imaging data to predict the cross-sectional area of the median nerve on ultrasound images at one, three, and six months after median nerve hydro-release treatment. The results showed that the multilayer perceptron model performed the best, with R^2 values of 0.906, 0.684, and 0.667 at one-, three- and six months post-treatment, respectively. They also reported that CTSI scores improved after the treatment, but they did not create a machine-learning model to predict the change in post-treatment CTSI. Harrison et al. [16] developed a binary classification model to predict whether the rate of Quick DASH improvement after CTR surgery would exceed the minimal clinically important difference (MCID), with predictive accuracies of 0.72 and 0.76 for improvements in functional and symptom scores after CTR surgery, respectively. In the study, they excluded items 7 and 8 among the Quick DASH disability items because those items are related to social activities. Itsubo et al. [17] compared the responsiveness of patients with the Quick DASH-JSSH and CTSI-JSSH three months after CTR and reported that CTSI-JSSH showed higher responsiveness than DASH-JSSH. Therefore, we selected CTSI-JSSH as a patient-oriented questionnaire. There are several reports on MCIDs for CTSI, with values ranging from 1.04~1.4 for CTSI-SS and 0.4~1.6 for CTSI-FS [18–20]. De Kleermaeker et al. [21] reported that MCID cutoff values were not constant and should be

determined based on individual preoperative CTSI scores. The change in CTSI-JSSH score in this study was 1.20 for CTSI-JSSH-SS and 0.99 for CTSI-JSSH-FS, which is comparable to the findings pf previous reports [18–21]. The RF model showed the highest score in the CTSI-JSSH-SS prediction model. The preoperative CTSI-JSSH score and age were the most important predictors. The RF model also showed the best performance in the prediction of CTSI-JSSH-FS. The important predictors in the RF model were the preoperative CTSI-JSSH score and age. Itsubo et al. [17] reported that improvements in the CTSI-JSSH score at six months postoperatively correlated with the CTSI-JSSH score itself, rather than with the preoperative electrophysiologic severity. Our results were consistent with their report. In this study, the R^2 values were 0.768 for CTSI-JSSH-SS and 0.846 for CTSI-JSSH-FS. Although the R^2 values are not as high as in other studies [15], this study is valuable as the first report of changes in CTSI-JSSH after surgery using machine learning.

To summarize the results of this study on CTSI-JSSH, the Random Forest model was the best predictor of postoperative change, with a lower preoperative severity score and functional score and younger age being important for predicting postoperative CTSI cores.

This is also the first study to use machine learning to predict the change in electrophysiological severity after CTR. The LGBM model showed the highest accuracy in the electrophysiological severity prediction model. Preoperative DML and age were the most important features for prediction. These have been reported as statistically significant factors that influence the change in electrophysiological severity after CTR surgery [18], and the machine learning predictions in this study showed similar trends. This result is similar to the items significant in the Pearson's correlation coefficient test.

The limitations of this study are as follows. First, since the study included patients who had undergone CTR, the number of electrophysiological mild cases (Stages 1 and 2) was small. Second, the number of explanatory factors was small: blood tests, BMI, and medical history were not included as explanatory factors.

5. Conclusions

Machine learning models were used to predict the changes in the electrophysiologic severity and CTSI-JSSH score one year after surgery. Important features for prediction were preoperative DML and age in the electrophysiological severity classification model, and preoperative CTSI-JSSH scores and age in the CTSI-JSSH-SS&FS prediction model.

Author Contributions: Conceptualization, A.I. and T.K.; methodology, A.I.; software, A.I. and F.T.; formal analysis, A.I. and F.T.; investigation, A.I.; resources, A.I., F.T. and T.K.; writing—original draft preparation, A.I. and S.L.; writing—review and editing, A.I., S.L. and T.K.; visualization, A.I.; supervision, S.L. and T.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: This study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Ethics Committee of Kobe Rosai Hospital (No. 4-04 approval date 9 August 2022).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available upon request from the corresponding author. These data are not publicly available because of confidentiality concerns.

Conflicts of Interest: Author Stefano Lucchina was employed by Locarno Hand Center. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Bickel, K.D. Carpal tunnel syndrome. *J. Hand. Surg. Am.* **2010**, *35*, 147–152. [CrossRef] [PubMed]
2. Fajardo, M.; Kim, S.H.; Szabo, R.M. Incidence of carpal tunnel release: Trends and implications within the United States ambulatory care setting. *J. Hand. Surg. Am.* **2012**, *37*, 1599–1605. [CrossRef]
3. Phalen, G.S. The carpal-tunnel syndrome. Clinical evaluation of 598 hands. *Clin. Orthop. Relat. Res.* **1972**, *83*, 29–40. [CrossRef]
4. Aversano, F.J.; Goldfarb, C.A.; Gelberman, R.H.; Calfee, R.P. Utility of the Carpal Tunnel Syndrome-6 for Predicting the Outcomes of Carpal Tunnel Release. *J. Hand. Surg. Am.* **2022**, *47*, 944–952. [CrossRef]
5. Pimentel, B.F.R.; Faloppa, F.; Tamaoki, M.J.S.; Belloti, J.C. Effectiveness of ultrasonography and nerve conduction studies in the diagnosing of carpal tunnel syndrome: Clinical trial on accuracy. *BMC Musculoskelet. Disord.* **2018**, *19*, 115. [CrossRef]
6. Bland, J.D. Do nerve conduction studies predict the outcome of carpal tunnel decompression? *Muscle Nerve* **2001**, *24*, 935–940. [CrossRef]
7. Kanatani, T.; Nagura, I.; Kurosaka, M.; Kokubu, T.; Sumi, M. Electrophysiological assessment of carpal tunnel syndrome in elderly patients: One-year follow-up study. *J. Hand. Surg. Am.* **2014**, *39*, 2188–2191. [CrossRef]
8. Peng, Z.; Wang, Y.; Wu, X.; Yang, S.; Du, X.; Xu, X.; Hu, C.; Liu, W.; Zhu, Y.; Dong, B.; et al. Identifying High Gleason Score Prostate Cancer by Prostate Fluid Metabolic Fingerprint-Based Multi-Modal Recognition. *Small Methods* **2024**, *8*, 2301684. [CrossRef]
9. Lyu, X.; Liu, J.; Gou, Y.; Sun, S.; Hao, J.; Cui, Y. Development and validation of a machine learning-based model of ischemic stroke risk in the Chinese elderly hypertensive population. *View* **2024**, *5*, 20240059. [CrossRef]
10. Imaeda, T.; Uchiyama, S.; Toh, S.; Wada, T.; Okinaga, S.; Sawaizumi, T.; Nishida, J.; Kusunose, K.; Omokawa, S. Clinical Outcomes Committee of the Japanese Orthopaedic A, the Impairment Evaluation Committee of the Japanese Society for Surgery of the H. Validation of the Japanese Society for Surgery of the Hand version of the Carpal Tunnel Syndrome Instrument. *J. Orthop. Sci.* **2007**, *12*, 14–21. [CrossRef]
11. Kanatani, T.; Harada, Y.; Nagura, I.; Takase, F.; Lucchina, S. Evaluation of carpal tunnel release outcomes by the Japanese version of the carpal tunnel syndrome questionnaire compared to an electrophysiological severity grade. *J. Orthop. Sci.* **2021**, *26*, 1004–1007. [CrossRef]
12. 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] [PubMed]
13. 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]
14. Shinohara, I.; Inui, A.; Mifune, Y.; Nishimoto, H.; Yamaura, K.; Mukohara, S.; Yoshikawa, T.; Kato, T.; Furukawa, T.; Hoshino, Y.; et al. Using deep learning for ultrasound images to diagnose carpal tunnel syndrome with high accuracy. *Ultrasound Med. Biol.* **2022**, *48*, 2052–2059. [CrossRef] [PubMed]
15. 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]
16. Harrison, C.J.; Geoghegan, L.; Sidey-Gibbons, C.J.; Stirling, P.H.; 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]
17. Itsubo, T.; Uchiyama, S.; Momose, T.; Yasutomi, T.; Imaeda, T.; Kato, H. Electrophysiological responsiveness and quality of life (QuickDASH, CTSI) evaluation of surgically treated carpal tunnel syndrome. *J. Orthop. Sci.* **2009**, *14*, 17–23. [CrossRef]
18. Özyürekoğlu, T.; McCabe, S.J.; Goldsmith, L.J.; LaJoie, A.S. The minimal clinically important difference of the Carpal Tunnel Syndrome Symptom Severity Scale. *J. Hand. Surg. Am.* **2006**, *31*, 733–738, discussion 739–740. [CrossRef]
19. Kim, J.K.; Jeon, S.H. Minimal clinically important differences in the Carpal Tunnel Questionnaire after carpal tunnel release. *J. Hand. Surg. Eur. Vol.* **2013**, *38*, 75–79. [CrossRef]
20. Ozer, K.; Malay, S.; Toker, S.; Chung, K.C. Minimal clinically important difference of carpal tunnel release in diabetic and nondiabetic patients. *Plast Reconstr. Surg.* **2013**, *131*, 1279–1285. [CrossRef]
21. De Kleermaeker, F.G.; Boogaarts, H.D.; Meulstee, J.; Verhagen, W.I. Minimal clinically important difference for the Boston Carpal Tunnel Questionnaire: New insights and review of literature. *J. Hand. Surg. Eur. Vol.* **2019**, *44*, 283–289. [CrossRef]

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.