Original Manuscript

Journal of Rehabilitation and Assistive Technologies Engineering
Volume 8: 1–11

DOI: 10.1177/20556683211044640
journals.sagepub.com/home/jrt

Classifying and tracking rehabilitation interventions through machine-learning algorithms in individuals with stroke

Victor C Espinoza Bernal[1], Shivayogi V Hiremath[1,2], Bethany Wolf[3], Brooke Riley[3], Rochelle J Mendonca[4] and Michelle J Johnson[5,6]

Abstract

Introduction: Stroke is the leading cause of disability worldwide. It has been well-documented that rehabilitation (rehab) therapy can aid in regaining health and function for individuals with stroke. Yet, tracking in-home rehab continues to be a challenge because of a lack of resources and population-scale demands. In order to address this gap, we implemented a methodology to classify and track rehab interventions in individuals with stroke.

Methods: We developed personalized classification algorithms, including neural network-based algorithms, to classify four rehab exercises performed by two individuals with stroke who were part of a week-long therapy camp in Jamaica, a low- and middle-income country. Accelerometry-based wearable sensors were placed on each upper and lower limb to collect movement data during therapy.

Results: The classification accuracy for traditional and neural network-based algorithms utilizing feature data (e.g., number of peaks) from the sensors ranged from 64 to 94%, respectively. In addition, the study proposes a new method to assess change in bilateral mobility over the camp duration.

Conclusion: The results of this pilot study indicate that personalized supervised learning algorithms can be used to classify and track rehab activities and functional outcomes in resource limited settings such as LMICs.

Keywords

Artificial neural networks, classification, machine-learning, stroke, global health, low–middle-income country (LMIC)

Date received: 27 February 2020; accepted: 20 August 2021

Introduction

Stroke is a leading cause of long-term disability for survivors.[1] Functional limitations, such as mobility impairments,[2,3] language deficits,[4,5] and paralysis, greatly reduce the quality of life.[1,6] Inpatient and outpatient rehabilitation (rehab) improves the health and function of individuals with stroke. However, availability of rehab services for individuals with stroke in low- and middle-income countries (LMICs) may be limited due to lack of resources, healthcare facilities being present mainly in larger cities, transportation barriers, and large population-scale demand.[7] One of the approaches to address this need may be to

[1]Department of Health and Rehabilitation Sciences, Temple University, Philadelphia, PA, USA
[2]Director of Personal Health Informatics & Rehabilitation Engineering Laboratory, Temple University, Philadelphia, PA, USA
[3]Friends of the Redeemer United, St Elizabeth, Jamaica
[4]Vagelos College of Physicians and Surgeons, Columbia University, New York, NY, USA
[5]Department of Physical Medicine and Rehabilitation, University of Pennsylvania, Philadelphia, PA, USA
[6]Director of Rehabilitation Robotics Lab, University of Pennsylvania, Philadelphia, PA, USA

Corresponding author:
Shivayogi V Hiremath, Personal Health Informatics & Rehabilitation Engineering (PHIRE) Laboratory, Temple University, Pearson Hall 40, 1800 North Broad Street, Philadelphia, PA 19121, USA.
Email: Shiv.Hiremath@temple.edu

use cost-effective wearable technology, which has the potential to assist clinicians in tracking and providing rehab interventions in the community.[8–10] However, a major limitation that needs to be addressed is the limited functional improvement at home due to *low and varying* patient adherence (20–35% overall) with prescribed exercises.[11–14] Research indicates that a technology coach, which supports patients around-the-clock, may assist patients with adherence.[15–17]

The applications of machine-learning algorithms are not new to stroke-related research.[18–21] Prior research by Manini et al. used a support vector machine (SVM) classifier algorithm to distinguish between gait in elderly patients and patients with stroke and Huntington's disease.[18] The SVM classier algorithm used a combination of time and frequency domain features with group-specific hidden Markov models (HMMs) to classify 90.5% of patients to the right group in a laboratory setting. Furthermore, the research suggested that investigating left-right symmetry through HMM can yield more accurate results given that gait is cyclical in nature. Xu et al. used a combined approach of using Dynamic Time Warping (DTW) algorithm and Naïve Bayes (NB) classifier to detect and classify walking compared to other activities.[21] According to their results, the combined method outperformed each technique individually in terms of accuracy. While the hybrid approach achieved a higher accuracy (97.3%), the DTW and NB methods achieved 93.4% and 93.7%, respectively. In another study, Roy *et al.* used a combination of neural networks and a neuro-fuzzy inference technique to classify activities of daily living in people with stroke.[22] The classification algorithms used data from surface electromyography and accelerometer sensors to classify 11 activities with an average accuracy of 95.0%. Roy *et al.* highlight that multi-layered neural networks were more efficient than single-layered ones.[22] Leveraging prior research, we are proposing the development of personalized machine-learning algorithms for individuals with stroke. Personalized machine-learning algorithms are defined as algorithms developed for each individual to better assess their limb movement in the presence of biomechanical variations due to physical impairments. These new techniques have the potential to improve classification of rehab interventions in individuals with stroke.

The primary objective of this pilot study was to develop and evaluate personalized machine-learning algorithms to detect and track rehab exercises for people with stroke. The personalized machine-learning algorithms developed included SVM, random forests (RFs), artificial neural networks (ANN) using stochastic gradient descent (SGD), and ANN using Adaptive moment parameter optimization. A secondary objective was to use the detected activities to assess change in bilateral mobility for upper and lower limbs during rehab exercises over time. We assessed bilateral upper or lower limb mobility of an individual with stroke by developing a new method called SymMetric, which provides a personalized value that captures change in mobility of paretic (affected) side as compared to non-paretic (non-affected) side of a person.

Materials and methods

Protocol

This study was conducted in Jamaica, an LMIC in the Caribbean, with a population of approximately 2.7 million people and a stroke prevalence 74.3 per 100,000.[23] Participants signed a written consent form prior to participating in the study. The study was considered a case study and approved by the Jamaican Government's Institutional Review Board. This pilot study included two cases studies, which comprised of data collected from two participants with stroke. Both participants participated in an intense and structured stroke rehab program involving upper and lower extremity functional activities for 7 h/day for a week in Jamaica.

Group mat activities included therapeutic exercises that focused on preventing contractures, improving postural control, building capacity in large muscle groups, and neuromuscular re-education. Furthermore, the rehab exercises were provided at *four treatment stations* that focused on a distinct aspect of motor function. The *gait station* included a custom body weight–supported treadmill training and over-ground gait training. Over-ground gait training was conducted both indoors and outdoors on a variety of surfaces with emphasis on normal gait kinematics. In the *balance station*, lower level patients worked on static and dynamic sitting balance, transfer training and static standing with manual assistance or a harness. Higher level patients worked on postural control during dynamic standing balance activities. In the *upper extremity station*, patients with lower level arm function focused on neuromuscular re-education with neuromuscular electrical stimulation and upper extremity exercises. Higher level patients worked on capacity building active exercises, purposeful movement with and without neuromuscular electrical stimulation, and gross and fine motor activities. In the *free station*, patients and therapists were able to focus on a deficit or functional activity that required more attention and repetition. In addition to the intervention stations, patients also participated in group activities that consisted of relay races, group volleyball, musical chairs, and dance contests. These activities made participation in rehabilitation fun and engaging.

Figure 1 shows an individual with stroke participating in a gait training exercise during the rehab program. The analysis in this study was based on data collected by wearable sensors, placed on participant's triceps and ankles bilaterally (Figure 1), for four of the 7 days when the participants performed rehab exercises in a semi-structured environment. Table 1 shows the four rehab exercises that participants performed, and their corresponding classification labels assigned. Over-ground gait training was conducted both indoors and outdoors on a variety of surfaces with emphasis on normal gait kinematics. On the first day, a clinician annotated the start and end times of each activity trial during data collection, which was used as the reference for developing and testing of the classification algorithms. This led to 180 min of annotated data for each participant (45 min per rehab exercise) for the first day. Due to the approximately equal duration of the rehab interventions, we obtained a balanced dataset for each activity. Additionally, the duration of data collected for the other 3 days was 600 and 690 min, respectively, for participants 1 and 2.

Instrumentation

The wearable device consists of a three axis accelerometer, a galvanic skin response sensor, a heat flux sensor, and a skin temperature sensor. For the purposes of this study, we only analyzed feature data extracted from the three axis accelerometer. In this study, we chose a tri-axial accelerometer-based sensor as it is embedded in a wearable device which is easy to use, and captures high-fidelity movement data of the person wearing it.[24,25] Additionally, gyroscope-based sensors consume several orders of magnitude more power than accelerometers.[25] Some of the disadvantages of using a wearable sensor include wearing the sensor and recharging the device on a regular basis.[24] The wearable device used developed a model, based on static and dynamic calibration tests, to measure three axis acceleration due to gravity and body movements. Static calibration tests provide a reference measurement of acceleration values for the tri-axial accelerometer.[26] The static test was conducted by the manufacturer through a precision inspection table using gravity as the reference. Dynamic calibrations tests provide the performance of the sensor through a range of frequencies and gravitational loads.[26] The dynamic calibration tests were conducted by the manufacturer through a vibration test system (i.e., a “shake table”). The device then removed the acceleration component due to gravity to store acceleration due to body movements. The three axis accelerometer sampled movement data at 32 Hz with a range of ± 39.2 m/s^2. Furthermore, the device used a low-pass moving average filter of 32 samples to obtain smoothed acceleration data in three axes, which were used to estimate feature data every minute. SenseWear software, developed by Bodymedia Inc (Pittsburgh, PA, USA), was used to retrieve the average acceleration and feature data recorded every minute from the wearable sensors for the duration of the study (4 days) (Table 2). The feature data comprised of the average (representing overall limb motion), the mean absolute difference (representing variability of limb motion), and the number of peaks (representing turning points of limb motion) in transverse, forward, and longitudinal directions. Overall, we obtained 36 features from four sensors (three features per axis for three axes accelerometer) that were normalized based on the maximum value for each

Figure 1. A participant with stroke participating in a gait exercise. Sensors are highlighted by rounded rectangles on the upper arm and ankles.

Table 1. Exercises performed during the rehab program and their corresponding labels for the study.

Activity	Label
Gait training	0
Balance training	1
Upper extremity therapeutic exercise	2
Overall function task training	3

Table 2. Feature data retrieved from wearable sensors for transverse, forward, and longitudinal directions.

Feature	Description
Average	Average acceleration (m/s^2)
MAD	Mean absolute deviation for acceleration (m/s^2)
Peaks	Number of peaks per minute for acceleration (peaks/min)

feature. Principal component analysis was used to better visualize the 36 features. The visualization indicated that the rehab activities were not easily separable on three key principal components obtained from the PCA. Therefore, we chose to use all 36 features to develop and evaluate the classification algorithms.

Data processing

Each participant's annotated data were divided into two sets: training (80%) and validation (20%). While the training data were used to develop classification models, the validation data were used to gauge the model's performance on data not used for model development. In addition, a five-fold cross-validation procedure was used during the model development on the training data. The algorithms were developed on a computer with Windows 10 Professional Operating System (System properties: Intel(R) Core(TM) i7-8850 central processing unit at 1.8 Giga Hertz; Random Access Memory: 16 gigabytes; 64-bit x 64-based processor). All classification algorithms' performance were evaluated in terms of classification accuracy. Additionally, the ANN classification algorithm's performance was evaluated in terms of accuracy and cross-entropy loss. The performance of the best classifier algorithm was based on a Kruskal–Wallis test, due to the non-normal distribution of accuracy, for the four classification algorithms over 100 iterations. Statistical analyses were performed using IBM SPSS Statistics software (ver. 25.0, Armonk, NY), with a statistical significance value set at an alpha level of 0.05. The following sections highlight the classification algorithms developed and the evaluation of change in bilateral mobility for upper and lower limbs during rehab over time.

Classification

We used TensorFlow[27] to develop two types of ANN-based classification algorithms including ANN with SGD optimization and ANN with Adaptive moment parameter optimization to classify the rehab exercises. Additionally, we used Scikit-learn[28] to develop SVMs and RF classification algorithms. The SVM classification algorithm used a linear Kernel, with a penalty parameter of 80, to transform the feature data into a higher dimension. The transformed data was then classified into rehab exercises via hyperplanes. On similar lines, an RF algorithm comprising of multiple decision trees was developed. Within each decision tree, there are binary choices where the tree splits into branches. The quality of each split was quantified by the Gini index, a generalization of binomial variance. The parameters we used for the RF algorithm included nine nodes with a minimum of two splits by internal node (maximum depth), 90 decision trees, and the Gini Index as the chosen impurity function. The final classification was a result of the majority of votes amongst all decision trees.

Our ANN model development included two sub processes. First was forward propagation, which made a prediction based on the input data and the model's current weight values. Second was backpropagation, which entailed updating the weight values within the model towards an error minimum. The ANN model's architecture had one input layer (normalized feature data), multiple hidden layers, and an output layer. The output layer consisted of four nodes corresponding to the four possible outcomes in our categorical classification problem. We used *one-hot encoded* classification, which mapped our targets and model's prediction to a binary vector containing four binary elements in which the only element set to one corresponds to one of the possible outcomes. Equation (1) shows the mathematical operation of a neuron, where x_i is an element of an input vector x, $w_i$is the ith weight value of a layer containing i neurons and b is a constant or bias. Equation (2) shows the function (φ) which was the rectified linear unit or activation function. Once the input traveled through all the hidden layers, the output nodes were activated by the softmax function (Equation (3)) in order to produce a distribution amongst the possible outcomes. K is the dimensionality of input vector z. The weight values were adjusted using two optimization techniques independently, SGD and Adaptive moment parameter optimization (Adam). Figure 2 illustrates the process we used to develop an ANN algorithm with Adam for this study.

$$\varphi = \sum_i (x_i w_i + b) \tag{1}$$

$$\varphi(\mathrm{x}) = \max(0, x) \tag{2}$$

$$\sigma(\mathrm{x})_i = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}} \tag{3}$$

In more detail, SGD is an optimization technique used to reduce the model's classification error based on the network's *gradient*, which is the derivative of the model's prediction error with respect to its current weights. In our work, the ANN with SGD model's error was quantified by binary cross-entropy loss function; a function which uses the model's prediction ($\hat{y}$) and actual (y) correct values (Equation (4)). Once the gradient, $\nabla_w J(w)$, was calculated

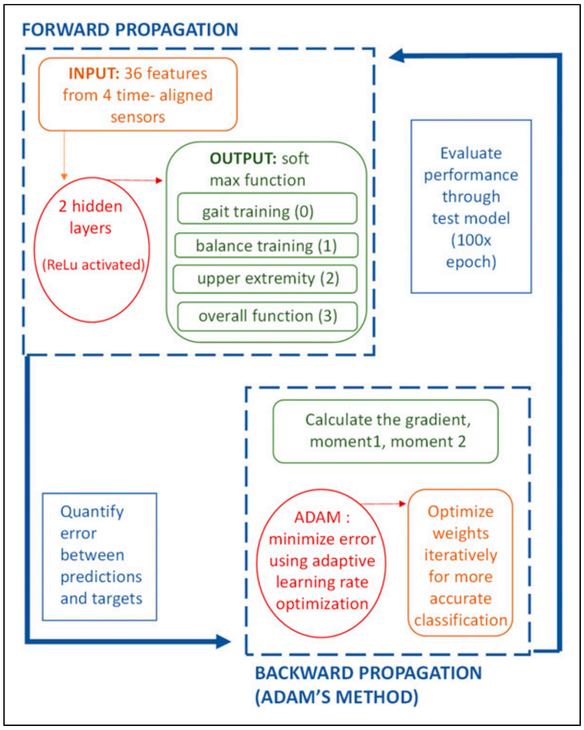

Figure 2. Activity Classification using ANN with Adam optimization.

(Equation (5)), the weight values in the model were iteratively optimized to minimize the error while using a constant hyper-parameter η (Equation (6)). Each iteration of the forward and backward propagation comprised an epoch.

$$L\left(y,\hat{y}\right) = -\sum_i y_i \log\left(\hat{y}_i\right) \tag{4}$$

$$\nabla_w J(w) = \frac{d\left(L\left(y,\hat{y}\right)\right)}{dw} \tag{5}$$

$$w = w - \eta \nabla_w J(w) \tag{6}$$

In addition to SGD, which is highly dependent on robust weight initialization and slow to reach loss convergence,[29] we have developed and evaluated ANN with Adam optimization. Adam is an augmented version of SGD which utilizes the mean (first moment) and variance (second moment) of the gradient in order to update the weights (Equation (7)). The $\hat{m}_t$ and $\hat{v}_t$ are the estimated moments which were initialized to zero (at t = 0) and computed over several iterations (Equations (8) and (9)). The β_1 and β_2 are coefficients to compensate for the zero initialization of the momentum and velocity. Adam optimizes faster than SGD by adapting the learning rate and using sparse features to obtain a faster convergence rate. In order to compensate for the bias towards zero, we computed bias as corrected estimates following Kingma and Ba's model.[30] Equations (10) and (11) represent this adjustment. The ANN algorithms were trained for 100 epochs (iterations of forward and backward propagation). Lastly, we recorded the algorithm's performance with all possible sensor permutations for each participant.

$$\mathrm{w_t} = \mathrm{w_{t-1}} - \frac{\alpha \hat{\mathrm{m}}_\mathrm{t}}{\sqrt{\hat{\mathrm{v}}_\mathrm{t}} + \eta}\hat{\mathrm{m}}_\mathrm{t} \tag{7}$$

$$m_t = \beta_1 m_{t-1} + (1-\beta_1)\nabla_w J(w) \tag{8}$$

$$v_t = \beta_2 v_{t-1} + (1-\beta_2)\nabla_w J^2(w) \tag{9}$$

$$\hat{m}_t = \frac{m_t}{1-\beta_1} \tag{10}$$

$$\hat{v}_t = \frac{v_t}{1-\beta_2} \tag{11}$$

The SymMetric method

Post-classification, we designed a new approach called SymMetric method which can quantify the relative limb functionality for the affected side as compared to the non-affected side in an individual with stroke. Prior research by Lang et al. compared the overall bilateral mobility of upper extremities with the ratio of mobility from affected and non-affected sides.[19] The resulting two-dimensional plot shows an individual's upper limb mobility,[19] but its clinical application is challenging due to personal interpretation. On the other hand, we combined both the frequency and magnitude of movement information from the affected and non-affected sides of an individual to a single value, which can be used to compare within-participant performance over days and activities.

The process we followed included using magnitude ratio and bilateral magnitude, two variables which were previously proposed.[19] Magnitude ratio is calculated as the natural logarithm (ln) of the vector magnitude of the affected limb divided by the magnitude of the non-affected limb vector magnitude (Equation (12)). Bilateral magnitude is calculated as the sum of the magnitude vectors of both limbs (Equation (13)). Figure 3 shows the bilateral magnitude and magnitude ratio for affected and non-affected sides of participant 1 with respect to the center for upper extremity therapeutic exercises during day 1.

$$\textit{magnitude ratio} = \ln\frac{L_a}{L_n} \tag{12}$$

$$\textit{bilateral magnitude} = \ ||L_n|| + ||L_n|| \tag{13}$$

First, we calculated the pair-wise difference of non-affected (A) and affected side (B) of mobility. Second, we calculated the average value by dividing the sum of difference with the number of pairs that had either one of both values for non-affected or affected sides, which provided us with the amount of symmetry present in an individual's bilateral upper or lower extremity movement. We then applied a hyperbolic tangent function to the average difference, which provided us with a single SymMetric value ranging from −1 to +1 (Equation (14)). This step provided us with three points of reference to interpret the symmetrical limb movement (Figure 4). The points of reference included: 1 for non-affected limb movement, 0 for symmetrical limb movement, and −1 for affected limb movement.

$$SymMetric\ value = tanh(AVG\ (A - B)) \tag{14}$$

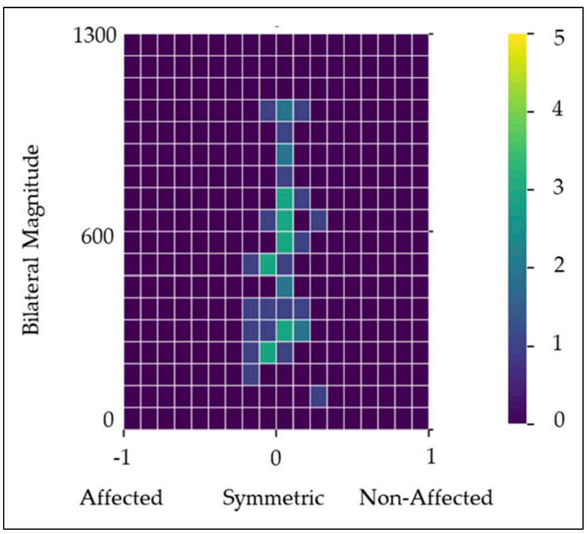

Figure 3. Affected and non-affected halves of a bivariate histogram. Bilateral magnitude (y-axis) versus magnitude ratio (x-axis) for affected (left side: 0 to −1) and non-affected (right side: 0 to 1) sides of participant 1 for upper extremity therapeutic exercises during Day 1.

In an individual with hemiparesis, we expect to see more activity on the non-affected side, which indicates that the SymMetric value will be positive rather than a negative number. Possible limitation of this approach is that the hyperbolic tangent function converges if the absolute value of the weighted average becomes greater than three.[31] This multi-angular analysis allows us to assess the mobility and change in mobility of the participant during a given activity over time.

Results

This pilot study included two male participants with stroke. Participant 1 had a stroke 31 months before the stroke camp, was 47 years old, and had left hemiplegia, which resulted in a severe upper limb impairment and moderate/mild lower limb impairment. Participant 2 had a stroke 3 months prior to the stroke camp, was 64 years old, and had right hemiplegia, which resulted in moderate upper limb impairment, and severe/moderate lower limb impairment. Table 3 shows the clinical assessments at the beginning and end of the rehab program. Although Participant 1 had lower scores for most clinical assessments, he was faster at the Box and Block test which led to a higher score on the right side. On the other hand, Participant 2 was slower at the Box and Block test but was able to perform the test with both of his upper extremities.

The average (SD) time it took to train SVM, RF, ANN with Adam, and ANN with SGD algorithms for 100 iterations (20*5-fold cross-validation) were 0.01 s (s) (0.00 s), 0.22 s (0.20 s), 21.80 s (20.64 s), and 22.60 s (20.10 s). These training times indicated that ANN-based algorithms were computationally more intensive than SVM and RF algorithms. Figures 5 and 6 shows the classification performance for four different personalized algorithms over 20 iterations of five-fold cross-validation for both participants. The highest median accuracy as well as the best performing

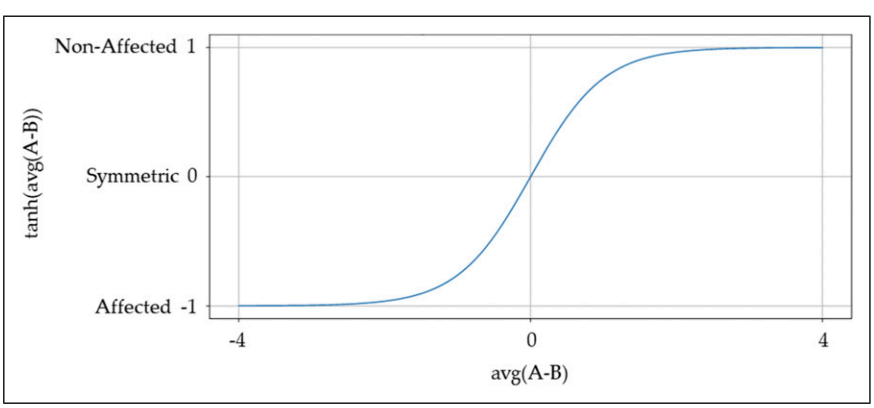

Figure 4. The SymMetric method shows a single value (y-axis) for average difference in frequencies of limb mobility for non-affected (A) and affected (B) sides (x-axis).

Table 3. Clinical assessments of participants at the beginning (Pre) and end (Post) of the rehab program.

	Participant 1		Participant 2	
Test	Pre	Post	Pre	Post
Box and blocks test (R – Right; L - left)	R: 59; L: 0	R: 65; L: 0	R: 24; L: 34	R: 24; L: 36
Fugl-Meyer (upper extremity motor; out of 66)	11	13	28	38
Berg balance test (out of 56)	26	27	51	53
6 minute walk test (feet)	300	400	790	1205
Timed up and go (sec)	32.3	29.2	24.6	11.4
10 meter walk test (sec)	23.8	21.4	16.4	8.6
Mobility aid	Quad cane		Single point cane	

interquartile range was achieved by the ANN with Adam optimization. Based on the median performance and statistical comparison of the classification algorithms (Table 4), we chose to develop and evaluate personalized ANN algorithms with Adam optimization (here onwards ANN with Adam) for a varying number of sensors used for classifying four rehab activities. Both cross-entropy loss and accuracy metrics were averaged for all possible permutations when using less than four sensors. These results are displayed in Table 5. The results show that an increase in the number of sensors led to an increase in the average accuracy and a decrease in average cross-entropy loss for training and validation data.

The ANN with Adam algorithm was then used to predict activity classification for three remaining days of rehab intervention. Figures 7 (participant 1) and 8 (participant 2) shows training data, which was collected during day 1, and the predicted rehab exercises for the testing data based on personalized models for the remaining 3 days of the intervention.

In order to visualize mobility differences across types of rehab exercises and time, we plotted the values computed by our new SymMetric method for upper extremity movement. Figures 9 and 10 show the SymMetric plots for Days 1–4 for both participants. The personalized model developed from Day 1 predicted the rehab activities for the remaining days (Days 2, 3, and 4). The closer the SymMetric value to zero, the more symmetric the movement of the participant's limbs. Overall, the plots show that rehab exercises helped: (1) improve bilateral mobility over time, (2) initially improve and then revert to the original value, and (3) initially worsen and then improve bilateral mobility.

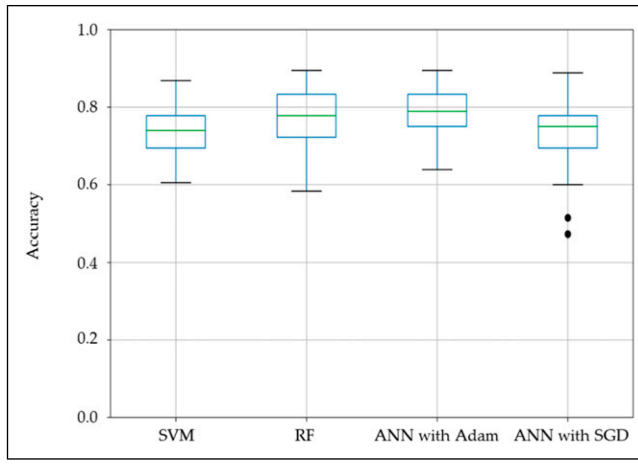

Figure 5. Performance of four classification algorithms for participant 1. The median (interquartile range) accuracy values obtained from 100 iterations (20*5-fold cross-validation) indicate that the ANN with Adam algorithm performed consistently well.

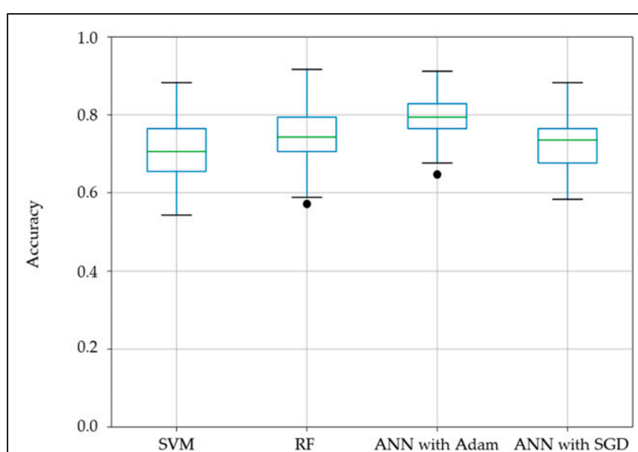

Figure 6. Performance of four classification algorithms for participant 2. The median (interquartile range) accuracy values obtained from 100 iterations (20*5-fold cross-validation) indicate that the ANN with Adam algorithm performed consistently well.

Discussion

In contrast to in-clinic rehab, recommendations for in-home exercises or therapy indicate that functional improvements experienced in a clinic do not continue at home, and in some cases, there is even a functional decline. To assist with tracking participants' rehab exercises in the community, we have (1) developed and evaluated personalized machine-

Table 4. Classification performance and comparison between each of the four personalized algorithms for participants 1 and 2.

		Participant 1	Participant 2
Classification algorithms accuracy (median (IQR))	SVM	0.74 (0.09)	0.71 (0.09)
	RF	0.78 (0.12)	0.75 (0.10)
	ANN with Adam	0.78 (0.09)	0.80 (0.07)
	ANN with SGD	0.75 (0.11)	0.75 (0.11)
Algorithms compared (*p*-value)	SVM–RF	$p < 0.001$[a]	$p < 0.001$
	SVM–ANN with Adam	$p < 0.001$[a]	$p < 0.001$[a]
	SVM – ANN with SGD	$p = 0.241$	$p = 0.016$[a]
	RF–ANN with Adam	$p = 0.202$	$p < 0.001$[a]
	RF–ANN with SGD	$p = 0.018$[a]	$p = 0.240$
	ANN with Adam–ANN with SGD	$p < 0.001$[a]	$p < 0.001$[a]

[a]indicates significantly different

Table 5. Average classification performance of ANN with Adam algorithm for participants 1 and 2.

		Average loss		Average accuracy	
	# of sensors used	Train	Validation	Train	Validation
Participant 1	4	0.20	0.59	0.94	0.82
	3	0.30	0.79	0.89	0.68
	2	0.71	0.92	0.73	0.63
	1	0.85	1.08	0.64	0.45
Participant 2	4	0.26	0.53	0.88	0.83
	3	0.31	0.72	0.89	0.76
	2	0.41	1.04	0.85	0.68
	1	0.77	0.92	0.70	0.63

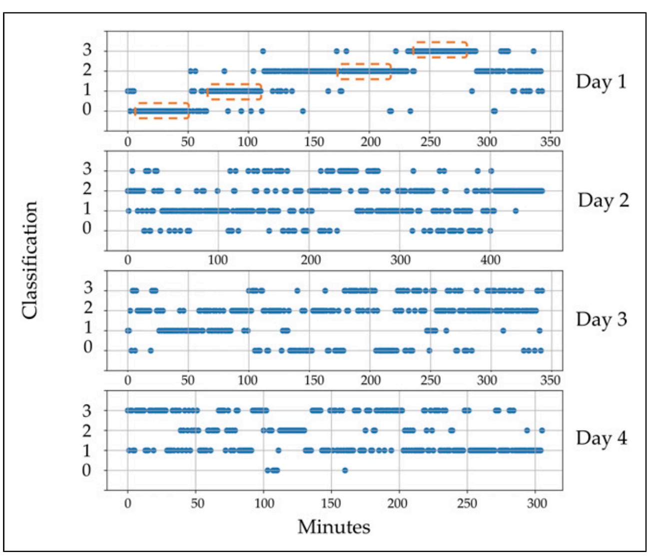

Figure 7. Rehab exercise predictions for participant 1. The x and y axes are time duration in minutes and rehab exercises, respectively. Classes 0, 1, 2, and 3 represent gait training, balance training, upper extremity therapeutic exercises, and overall function task training, respectively. While Day 1's predictions are based on the annotated data (highlighted with a dashed rounded rectangle box in orange), the rehab exercises predicted for Days 2, 3, and 4 are based on the personalized classification models developed from Day 1.

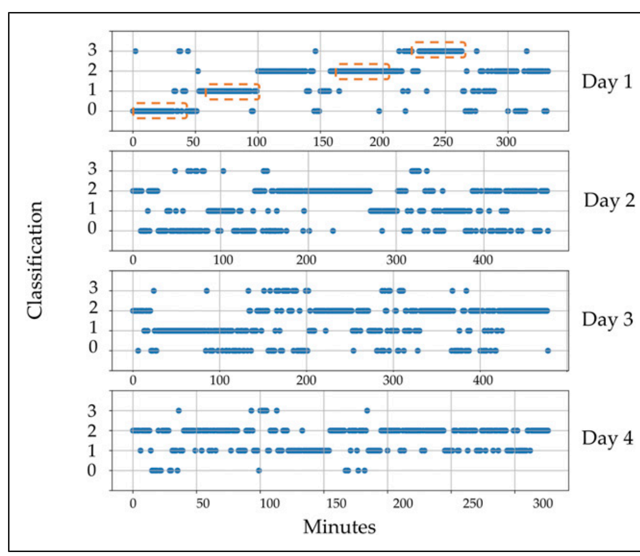

Figure 8. Rehab exercise predictions for participant 2. The x and y axes are time duration in minutes and rehab exercises, respectively. Classes 0, 1, 2, and 3 represent gait training, balance training, upper extremity therapeutic exercises, and overall function task training, respectively. While Day 1's predictions are based on the annotated data (highlighted with a dashed rounded rectangle box in orange), the rehab exercises predicted for Days 2, 3, and 4 are based on the personalized classification models developed from Day 1.

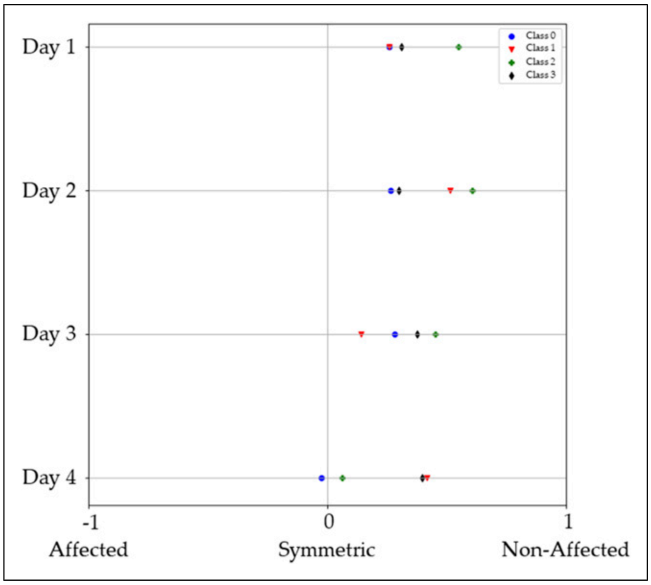

Figure 9. Bilateral mobility for participant 1 during the rehab program. The x and y axes are SymMetric bilateral movement and days of rehab exercises, respectively. Classes 0, 1, 2, and 3 represent gait training, balance training, upper extremity therapeutic exercises, and overall function task training, respectively. The plot shows that the change in bilateral upper limb movement over days for each participant varies based on their functional abilities. The closer the SymMetric value to zero, the more symmetric the movement of the participant's limbs. For example, in this participant we observed that: (1) upper extremity therapeutic exercises led to a symmetric use of bilateral upper limbs over time (Day 2 to Day 4); (2) balance training improved from Day 2 to Day 3, but went back to a value similar to Day 2; (3) overall function task showed a decrease in bilateral mobility from Day 2 to Day 3, and improved on Day 4; and (4) gait training showed a decrease in bilateral mobility from Day 2 to Day 3, with no prediction or value for Day 4 as the rehab therapy probably focused more on the other rehab activities.

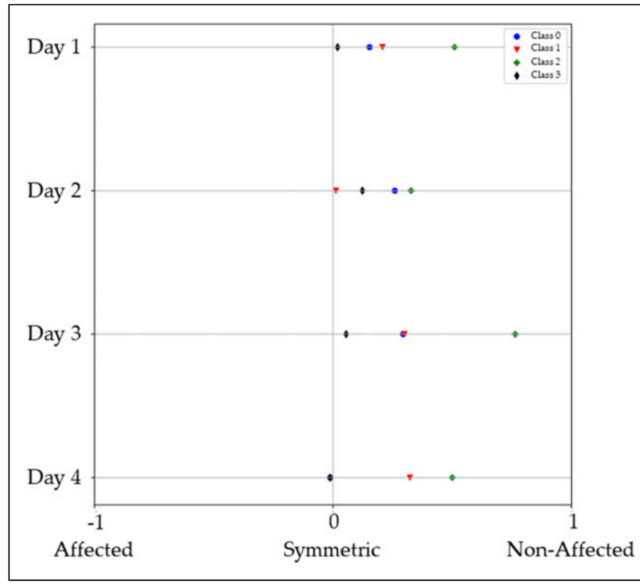

Figure 10. Bilateral mobility for participant 2 during the rehab program. The x and y axes are SymMetric bilateral movement and days of rehab exercises, respectively. Classes 0, 1, 2, and 3 represent gait training, balance training, upper extremity therapeutic exercises, and overall function task training, respectively. The plot shows that the change in bilateral upper limb movement over days for each participant varies based on their functional abilities.

learning algorithms to detect and track rehab exercises for people with stroke, and (2) developed a new methodology that tracks personalized rehab activities and estimates the change in bilateral mobility for upper and lower limbs during rehab exercises over time.

Personalized machine-learning algorithms

The results of this pilot study indicate that ANN models with Adam optimization achieved greater than 80% classification accuracy with low cross-entropy loss (0.6) for both participants (Table 5). In addition, the interquartile range from the 100 iterations (20 times of a 5-fold cross-validation, Figures 5 and 6) indicated other models achieved a similar range of classification accuracy. While the SVM classification algorithm achieved the lowest accuracy for both participants, the ANN with Adam achieved the highest accuracy. Furthermore, the classification accuracy increased with the number of sensors. In contrast, the model's loss rate increased as the number of sensors decreased.

The personalized algorithms were then used to assess which activities were performed by the individuals on the remaining days. Figures 7 and 8 show the rehab exercises that were predicted throughout the 4-day intervention of the rehab therapy program for both participants. The model performs well on the first-day data since it mainly consists of labeled data, while some patterns of the 45-min intervals can be observed on the second and third day. The predictions on last day data show small amount of gait-related rehab exercises as the therapy exercises focused more on a combination of balance training and upper extremity therapeutic exercise.

Change in bilateral mobility

Furthermore, the results of this pilot study indicate that the SymMetric method proposed by our group has potential to track changes in bilateral mobility over time and across various types of rehab exercises. Previous work by Lang et al.[19] suggested that a person without a hemiplegia will have a symmetric heat map indicating bilateral mobility. On the other hand, an individual with hemiplegia will have a pattern that is skewed to the non-affected side. While our current study was limited to two participants with stroke, the results suggested that SymMetric method can be used to quantify an individual with stroke's upper or lower extremity movement into a single value. In addition, the method focused on tracking change of affected limb's mobility compared to the non-affected side within rehab exercises and the intense rehab therapy program over time. Clinicians can use this quantitative measure to identify rehab activities that lead to symmetric limb movement for each individual leading to a personalized treatment plan.

The bilateral mobility scores for specific rehab activities and days indicates the need for personalized algorithms. This is further supported by the clinical measure differences (Fugl-Meyer scores) between participants. The clinical measures suggest that participant 2 would need less assistance than participant 1 in performing the rehab exercises and would likely demonstrate more independence with functional mobility in the community as well. Future research should evaluate how bilateral mobility during rehab exercises promoted improved daily use of the extremities in the home and the community. Furthermore, while both participants had upper extremity motor function improvement (Fugl-Meyer scores), only participant 2 showed an increase in mobility for the affected arm (Box and Blocks test). These results suggest that the clinical measures used in this study were not sensitive to small bilateral changes. Future studies should use clinical measures such as the Upper Extremity Motor Activity Log that capture both the amount and quality of arm use in individuals with stroke.[32] In addition, there is a need for developing clinical measures that capture change in bilateral mobility over time and relate it to kinematic information obtained from wearable devices.

Limitations

A limitation of this pilot study includes developing personalized algorithms and evaluating SymMetric method in just two individuals with stroke. Based on the results of this work, we will conduct a human subject research study in the community with a much larger sample size, as determined by power analysis. Furthermore, evaluating these models in a larger sample of participants will allow researchers to study variation in models' parameters for different subjects. While the current study indicated the performance of the personalized algorithms by presenting accuracy and loss entropy (ANN algorithm), future studies should assess specificity and sensitivity of the methodology in a large sample of participants. Another limitation of the study was that the clinicians were able to annotate the rehab activities and sensor data for the first day only due to lack of time and lack of staff resources during the rehab camp. Future studies should annotate the rehab activities and sensors data for multiple days, which would allow researchers to develop and train classification models for some of these days (training data) and gauge the model's performance on the other days not used for training (validation data).

Conclusions

Our work represents an attempt to automate the detection and quantification of change in limb mobility during rehab therapy for individuals with stroke. Personalized machine-learning algorithms combined with the SymMetric method for data obtained from wearable sensors has the potential to track rehab exercises in individuals with stroke in a semi-structured environment. Functional differences between participants with stroke can be captured through personalized algorithms, which may lead to tailored rehab exercises that promote bilateral and symmetric mobility. The work confirms that it is applicable to resource limited settings in high-income countries and to LMICs.

Acknowledgments

We would like to thank the students and staff at Friends of the Redeemer United, Department of Health and Rehabilitation Sciences at Temple University, and Department of Physical Medicine and Rehabilitation at the University of Pennsylvania for their support with this research study. We also thank Micaela Robalino and Alexandra Canori for their help with editing this manuscript.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Contributorship

SVH and MJJ conceptualized the study. MJJ, SVH, and VCEB developed the methodology. VCEB developed the software for analysis. VCEB and SVH conducted the formal analysis. MJJ, BW, BR, and RJM collected the data. SVH, MJJ, BW, and BR compiled the resources. BW, MJJ, SVH, and VCEB curated the data. VCEB, SVH, and MJJ wrote the first draft of the manuscript. All authors reviewed, edited, and approved the final version of the manuscript.

ORCID iDs

Shivayogi V Hiremath https://orcid.org/0000-0002-9708-1411
Michelle J Johnson https://orcid.org/0000-0002-7964-0304

References

1. Duncan PW, Zorowitz R, Bates B, et al. Management of adult stroke rehabilitation care: a clinical practice guideline. *Stroke* 2005; 36: e100–e143.
2. Collen FM, Wade DT and Bradshaw CM. Mobility after stroke: reliability of measures of impairment and disability. *Int Disabil Stud* 1990; 12: 6–9.

3. Harris JE, Eng JJ, Marigold DS, et al. Relationship of balance and mobility to fall incidence in people with chronic stroke. *Phys Ther* 2005; 85: 150–158.
4. Srikanth VK, Thrift AG, Saling MM, et al. Increased risk of cognitive impairment 3 months after mild to moderate first-ever stroke. *Stroke* 2003; 34: 1136–1143.
5. Basilakos A, Rorden C, Bonilha L, et al. Patterns of post-stroke brain damage that predict speech production errors in apraxia of speech and aphasia dissociate. *Stroke* 2015; 46: 1561–1566.
6. King RB. Quality of life after stroke. *Stroke* 1996; 27: 1467–1472.
7. Mukherjee D. and Patil CG. Epidemiology and the global burden of stroke. *World Neurosurg* 2011; 76: S85–S90.
8. Rodgers MM, Alon G, Pai VM, et al. Wearable technologies for active living and rehabilitation: current research challenges and future opportunities. *J rehabilitation assistive Tech Eng* 2019; 6: 2055668319839607.
9. Patel S, Hughes R, Hester T, et al. A novel approach to monitor rehabilitation outcomes in stroke survivors using wearable technology. *Proc IEEE* 2010; 98: 450–461.
10. Patel S, Park H, Bonato P, et al. A review of wearable sensors and systems with application in rehabilitation. *J NeuroEngineering Rehabil* 2012; 9: 21.
11. Sluijs EM, Kok GJ and Van der Zee J. Correlates of exercise compliance in physical therapy. *Phys Ther* 1993; 73: 771–782.
12. Kåringen I, Dysvik E and Furnes B. The elderly stroke patient's long-term adherence to physiotherapy home exercises. *Adv Physiother* 2011; 13: 145–152.
13. Miller KK, Porter RE, DeBaun-Sprague E, et al. Exercise after stroke: patient adherence and beliefs after discharge from rehabilitation. *Top Stroke Rehabil* 2017; 24: 142–148.
14. Yao M, Chen J, Jing J, et al. Defining the rehabilitation adherence curve and adherence phases of stroke patients: an observational study. *Patient Preference and Adherence* 2017; 11: 1435–1441.
15. López-Nava IH and Muñoz-Meléndez A. Wearable inertial sensors for human motion analysis: a review. *IEEE Sensors J* 2016; 16: 7821–7834.
16. Taylor PE, Almeida GJM, Kanade T, et al. Classifying human motion quality for knee osteoarthritis using accelerometers. In: International Conference of the Engineering in Medicine and Biology Society, 2010, pp. 339–343.
17. Huang K, Sparto PJ, Kiesler S, et al. A technology probe of wearable in-home computer-assisted physical therapy. In: Proceedings of the 32nd annual ACM Conference on Human Factors in Computing Systems. ACM, 2014, pp. 2541–2550.
18. Mannini A, Trojaniello D., Cereatti A, et al. A machine learning framework for Gait classification using inertial sensors: application to elderly, post-stroke and Huntington's disease patients. *Sensors* 2016; 16: 134.
19. Lang CE, Waddell KJ, Klaesner JW, et al. A method for quantifying upper limb performance in daily life using accelerometers. *J Vis Exp Jove*.
20. Dobkin BH, Xu X, Batalin M, et al. Reliability and validity of bilateral ankle accelerometer algorithms for activity recognition and walking speed after stroke. *Stroke* 2011; 42: 2246–2250.
21. Xu X, Batalin MA, Kaiser WJ, et al. Robust hierarchical system for classification of complex human mobility characteristics in the presence of neurological disorders. In: International Conference on Body Sensor Networks. IEEE, 2011, pp. 65–70.
22. Roy SH, Cheng MS, Chang S-S, et al. A combined sEMG and accelerometer system for monitoring functional activity in stroke. *IEEE Trans Neural Syst Rehabil Eng* 2009; 17: 585–594.
23. Ferguson T, Younger NO, Morgan ND, et al. Self-reported prevalence of heart attacks and strokes in Jamaica: a cross-sectional study. The Jamaica health and Lifestyle survey 2007–2008. *Res Rep Clin Cardiol* 2010; 1: 23–31.
24. Bergmann JHM and McGregor AH. Body-worn sensor design: what do patients and clinicians want?. *Ann Biomed Eng* 2011; 39: 2299–2312.
25. Twomey N, Diethe T, Fafoutis X, et al. A comprehensive study of activity recognition using accelerometers. *Informatics* 2018; 5: 27.
26. Gilbert BK, Haider CR, Schwab DJ, et al. A measurement-quality body-worn sensor-agnostic physiological monitor for biomedical applications. *Am J Biomed Eng* 2015; 5: 34–66.
27. Abadi M, Barham P, Chen J, et al. Tensorflow: a system for large-scale machine learning. In: 12th USENIX Symposium on Operating Systems Design and Implementation 2016. Savannah, GA, 2–4 November 2016.
28. Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: machine learning in python. *J Mach Learn Res* 2011; 12: 2825–2830.
29. Ruder S. An overview of gradient descent optimization algorithms. arXiv Prepr arXiv160904747.
30. Kingma DP and Ba J. Adam: a method for stochastic optimization. arXiv Prepr arXiv14126980.
31. Spanier J and Oldham KB. The hyperbolic tangent tanh (x) and cotangent coth (x) functions. In: *An atlas of functions*. Hemisphere, 1987, pp. 279–284.
32. Uswatte G, Taub E, Morris D, et al. The motor activity Log-28: assessing daily use of the hemiparetic arm after stroke. *Neurology* 2006; 67: 1189–1194, LP – 1194.