Machine Learning to Estimate the Amount of Training to Learn a Motor Skill

Moisés R. Santos[1], Eduardo D. F. Souza[1], Mateus B. F. Carvalho[1],
Alexandre C. M. Oliveira[1], Areolino de Almeida Neto[1], Marco R. Curado[2],
and Paulo R. A. Ribeiro[1(✉)]

[1] Universidade Federal do Maranhão, São Luís, Maranhão, Brazil
`moisesrs@lacmor.ufma.br, eduardo.dorneles.fs@gmail.com,`
`mateusbfrota@gmail.com, {alexandre.cesar,areolino}@ufma.br,`
`paulo.ribeiro@ecp.ufma.br`
[2] AbbVie Deutschland GmbH & Co. KG, Knollstrasse, Ludwigshafen, Germany
`marco.curado@abbvie.com`

Abstract. Machine Learning (ML) has been widely and successfully employed in different fields to estimate information from datasets. However, the necessary time to learn a motor task or to rehabilitate is mainly determined by the professional experience of medical doctor, physiotherapist and so on. Thus, this work introduces a software to measure the performance of subjects on a experiment performing a tracing task, which requires motor learning, and uses ML algorithms on the dataset acquired during this experiment. The task is divided into 1 session that has 3 blocks and each block is composed of 10 trials whereas each trial is one word. Furthermore, ML algorithms - namely k-nearest neighbours, decision tree, support vector machines and multilayer-perceptron neural network - are applied on the collected data from the experiment to estimate which block the subject currently is. The results demonstrated that there was motor learning, as well as that is possible to apply classification models to predict the block of the subject with decision tree achieving statistically significant (p-value < 0.01) best predictions. The proposed approach may be useful for health professionals when estimating the amount of training a patient requires to learn a motor task or rehabilitate.

Keywords: Motor learning · Rehabilitation · Machine Learning

1 Introduction

An enormous part of the Brazilian population has some kind of motor disease. According to the Brazilian Institute of Geography Statistics, its 2010 national census [18], 7% of the Brazilian population has some kind of motor disease,

e.g. stroke, cerebral palsy and congenital disorders. Then, neuroscience studies emphasizing motor learning and rehabilitation may pave the way to reveal solutions for these patients.

Some aspects of how motor learning may occur have been studied, e.g. [24] has a review of some relevant works, presenting the main components of motor learning: information extraction, classes of control, decisions and strategies. Furthermore, [24] also presents some processes of motor learning: error-based learning, reinforcement learning and use-dependent learning.

During motor rehabilitation sessions, the professional experience - from medical doctor, physiotherapist and so on - is the predominant information about the necessary time that a patient may require to rehabilitate or improve his/her performance. The absence of more details and information may demotivate the patient and create false expectations about his/her own progress.

Machine Learning (ML) can estimate patient's data from a historical dataset [8,17,20]. These ML algorithms have been widely employed in different fields such as health, education, economy, science, technology and logistic [5]. Furthermore, these algorithms might be used to predict the necessary time that a patient may require to learn a motor task or rehabilitate.

Thus, in addition to the professional experience, other information may be used to infer patient improvement, for example historical data from rehabilitation sessions of others patients. Then, Machine Learning algorithms may be convenient for this problem once they can be used to make predictions and estimate patients improvements from a historical dataset.

This paper presents a software tool to measure and compute the performance of subjects in a tracing task [12], as well as it compares different classifiers to estimate the amount of training to learn a motor skill.

This work is divided into five sections. Section 2, Motor Learning and Models of Prediction, details some aspects of Motor Learning as well as presents some works on Models of Prediction Applied to Motor Learning and Rehabilitation. Section 3, Material and Methods, explains how the experiment was conducted. Section 4, Results, presents and discusses the experiment assessment. Finally, Sect. 5, Conclusion, summarizes the work and presents some future works.

2 Motor Learning and Models of Prediction

2.1 Motor Learning

The main studies on motor learning present three components: information extraction, classes of control, decisions and strategies [24]. These components define how motor learning may occur. Information extraction is responsible for the capabilities that the motor learning system has to influence sensory organs [15]. Another interesting aspect of information extraction is how the brain filters information. People are only able to realize changes in the environment, while performing some motor task, when the change is related to the task in execution [22]. As a result, the need to extract information or process it only happens when it is relevant to execute actions of interest to the motor task.

Any kind of motor task demands decision making from extracted information, strategies are formulated to finish the task successfully considering how or when the movement must be performed [24]. In decision making experiments a decision variable is used to represent: the experience accumulation, an evidence and a value that must be interpreted by the decision rule to finally generate a choice for the participant [4]. A decision maker is interest in obtain positive feedback while avoids negatives. Positive feedback can be understood as correct decisions while negative feedback are errors, waste of time or efforts.

Classes of control are related to the means of how optimize the motor performance, they are subdivided in predictive control, reactive control, and biomechanical control [24]. Predictive control is associated to the sensorimotor outcome delays, the prediction is kept based on the individual's experience. Meanwhile, the reactive control is related to the updates in motor commands during the execution course after new sensory perception. Finally, the biomechanical control is related to the limbs movements that have a considerable factor in motor improvement. All the three classes are adaptable and contribute to the motor learning, depending on the task that has been performed is possible that some classes contribute in a more significant manner than others.

The processes of motor learning may indicate how motor learning is acquired in neurons while focus in the used information, they are: error-based learning, reinforcement learning and use-dependent learning [24]. The sensorimotor is capable to realize if the movement outcome has been performed as expected in the error-based learning. Sensory prediction errors are interesting because they show the causes for the failure. Reinforcement learning, differently from the error-based learning, takes positives and negatives values for each performed action, without any information regarding how the motor commands should be changed for a better performance. Reinforcement learning offers the advantage of contributing more to skill learning acquisition because of its rewards signals [1]. However, reinforcement learning mechanisms are not well understood yet and a better comprehension of what might be considered a reward to the motor system still exists [24]. Use-dependent learning refers to the capability of learning based on pure movements repetition even when there is no outcome information available [24].

2.2 Models of Prediction Applied to Motor Learning and Rehabilitation

A prediction model to evaluate stroke progression using Functional Independence Measure (FIM) score has been used and collected data from a motor rehabilitation center in Japan. A prediction model was built with Stepwise Multiple Regression and obtained a 0.68 correlation coefficient [23].

Another work that intended to predict FIM has used some characteristics from 87 patients in rehabilitation after stroke. In this case, a multiple linear regression has been adopted and obtained a 0.88 correlation coefficient [20].

A prediction model to estimate the level of independence in routine activities after stroke was created. A dataset with characteristics of 65 patients - with parameters such as age, time since stroke, functional independence measure

motor subscale score, functional independence measure cognitive subscale score, stroke impairment assessment set score, Berg balance scale score, and vitality index - was used to predict the future functional independence measure motor subscale score. They used Multiple Stepwise Regression to obtain the model and had a coefficient of determination computed of 0.60 [17].

Rather than using 3 level resolution per item on the Fugl-Meyer Assessment (FMA), [6] proposed another approach using 14 levels, i.e. obtaining more resolution, aiming to detect patient's smaller improvements. The proposed was to use 14 fractional digits resolution instead of 3, in 6 upper-extremity FMA items, so the patient's minor improvements could be more accurately detected. Some predict models were used - such as Linear Regression, Bayesian Linear Regression, Neural Network Regression, Boosted Decision Tree Regression and Decision Forest Regression - to convert these 3 levels to the 14 levels that were proposed. The best result was with a Neural Network Regression that obtained a 0.58 Pearson Correlation Coefficient.

Support Vector Machine has been used to predict the patients functional damage with stroke and achieved a 0.85 of accuracy [3]. This work shows how Machine Learning approaches may be used in rehabilitation scenarios.

Another interesting approach evaluates the applications of collected data from inertial wearable sensors to support decision making in rehabilitation clinics [21]. The collected data were used to predict the FIM from patients in rehabilitation. Machine learning algorithms were trained with 20 patients data in order to predict their FIM. The methods used were: Support Vector Regression (SVR), Linear Regression and Random Regression Forest. The algorithm with the best result was SVR with a 0.97 Pearson Correlation Coefficient.

Machine Learning was also employed to predict improvements of 55 patients in first stage treatment after stroke. The attribute to compute patient improvement was Motor (FIM) and the machine learning algorithm applied was the Support Vector Regression - which had three different kernels i.e. Polynomial, Gaussian and Linear - with a Pearson Correlation Coefficient between 0.71 and 0.81 [16].

3   Materials and Methods

3.1   Participants

The participants of this experiment were students at Federal University of Maranhão (UFMA). Initially a group of 3 healthy people, i.e. people without motor constraint, were volunteers in this experiment. All participants provided written authorization in order for their data to be used in a publication whereas at the same time this work maintains the confidentiality of the participants.

For the selection of participants, a neurological history was also carried out. They were asked about any constraint to perform a motor task. Subjects with any motor problem would be excluded from this study. This questionnaire is used to know which hand is used in daily living activities. The subject must use his/her

non-dominant hand, which in this case was the left hand and was measured with the "Edinburgh Handedness Inventory" [2], once only right-handed participants were considered in this study. The use of the non-dominant hand created the environment to learn a motor skill necessary to accomplish the task.

3.2   Study Design

A software for the Tracing Task, which was a task proposed by [12], was developed. The experiment replicated the Tracing Task once it is an interesting task to study motor learn. The software was developed using the Pygame Framework [19] of the Python programming language [14].

The task consists of subjects drawing a series of words with the aid of a Wacom tablet with their left hand, i.e. their non-dominant hand. The Tracing Task is depicted on Fig. 1.

Fig. 1. Tracing Task. The participant traces a word that is shown on the screen whereas the cursor on it is controlled with a pen on a tablet. Image Source: Own Authors

Each subject experiences one session, which has 3 blocks and each block is composed of 10 trials whereas each trial is one word. All the subjects performed the same words. The words used were chosen from a free database (Invoke IT Limited[1]). They are the most common words of 5 letters of the Brazilian Portuguese language. The words were written with a free font (League Script[2]).

The template and the traced word undergo a process of blurring so that participants who do not perfectly match the shape of the drawing are not harmed.

---

[1] https://invokeit.wordpress.com/frequency-word-lists/.
[2] https://www.theleagueofmoveabletype.com/.

The blurring process consists of a convolution of each image with a Gaussian kernel of size 50 × 50 pixels. These filtered images are thresholded, i.e. they are converted to binary images, and later a metric (percentage error), which is a difference between these binary images, is generated. This whole process is illustrated on Fig. 2.
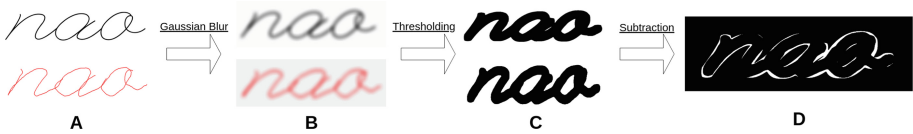

Fig. 2. Method to calculate the error in the Tracing Task. **(A)** is an example of a trial, the image above is the template image whereas the image below is the traced image. **(B)** presents the filtered images. **(C)** is the thresholding of the images. **(D)** shows the subtraction of the images. Percentage error is obtained from **D**. Image Source: Own Authors

After the data was acquired, they were formatted in Comma Separated Value (CSV) so that processing and analysis can be done. The CSV format was chosen for simplicity and for being widely used in statistical analysis.

3.3   Complementary Samples

Complementary samples were created from those obtained by the experiment in order to verify how the result would may be with more subjects. These samples were related to the average performance of the 3 blocks, for the two measures of performance: necessary time to trace the word and the percentage error.

A Gaussian noise with zero mean and standard deviation 0.01 (value defined based on the magnitude of the data) was added to the real data. A new dataset was generated, which contained 18 participants generated based on the 3 real subjects, i.e. for each one of these 3 participants 6 subjects were generated.

3.4   Classification Models

A subject performs a training, which is composed of session, blocks, and trials. ML algorithms are used - based on error and the time to perform the task - to predict the amount of blocks that the subject has performed. Classification models were built using the following algorithms: k-nearest neighbours (KNN), decision tree (AD), support vector machines (SVM) and multilayer-perceptron neural network (MLP) [9,10].

The scikit-learn [11] was used: (i) the K value in KNN was set to 3 and with Euclidian distance; (ii) the default configurations of decision tree were used; (iii) for SVM, the radial basis kernel was used; and (iv) MLP neural network used the logistic sigmoid function activation, the hidden layer size was set to 10, the max number of epochs was 10000 and the solver function chosen was Quasi-Newton method.

3.5   Validation

The validation method for the constructed models was 5-fold cross-validation whereas accuracy, recall, and specificity were used as evaluation metrics.

For statistical analysis the models were executed 10 times by randomly blending the dataset. An Analysis of Variance (ANOVA) was carried out to determine whether there was a statistically significant difference between the classifiers. The statistical analyzes were carried out using the R programming language. All statistical tests were two-sided and a p-value < 0.01 was used as statistically significant.

4   Results

4.1   Motor Learning

The results shown on this subsection are from the 3 real subjects. The motor learning characteristic can be observed on Fig. 3. According to [7] and [13], the motor learning occurs when there is a balance between speed and accuracy in performing a task, as depicted on Fig. 3.
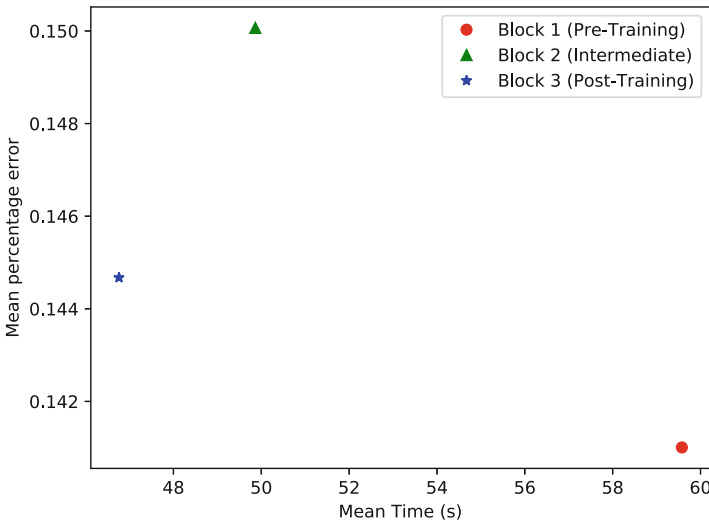

Fig. 3. Motor Learning. The red dot represents the average performance of participants in Block 1 (Pre-Training). The green triangle represents the average performance of participants in Block 2 (Intermediate). The blue star represents the average performance of participants in Block 3 (Post-Training). The mean performance is a function of the mean time of movement (abscissa) and the mean percentage error per block (ordinate). Image Source: Own Authors (Color figure online)

On Fig. 3, it can be seen that there is a motor learning comparing Pre-Training with Post-Training. In Pre-Training, the participants had on average a

small percentage error but the average execution time was greater in comparison to the other blocks. In the Intermediate training, the participants decreased the average time but increased the average percentage error. In post-training it is possible to observe a balance between time and average percentage error. Thus, there was motor learning between Block 1 and Block 3.

4.2   Comparative Analysis of Models

The results shown on this subsection are from the 18 simulated subjects. Figures 4, 5 and 6 show the performance for three different metrics - Accuracy, Specificity and Recall, respectively - for different classification models: Neural Network Multi Layer Perceptron (MLP); Support Vector Machine (SVM) with kernel RBF; K-Nearest Neighbors (KNN); and Decision Tree (AD).
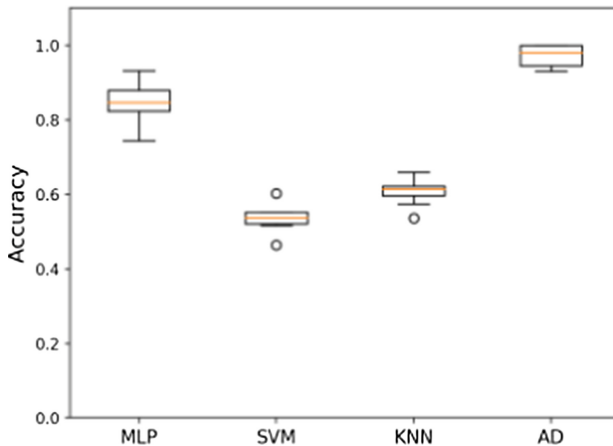

Fig. 4. Box plot of the performance for different classification models and the metric is Accuracy. The models are Neural Network Multi Layer Perceptron (MLP); Support Vector Machine (SVM) with kernel RBF; K-Nearest Neighbors (KNN); and Decision Tree (AD). Image Source: Own Authors

The Decision tree algorithm presented a superior performance compared to the others with: an average accuracy of 0.97; specificity (0.98); and recall (0.95). The mean performance of the neural network MLP was: accuracy = 0.84, specificity = 0.88 and recall 0.78. The performance of the SVM method was: accuracy = 0.53, specificity = 0.55 and recall 0.55. The performance of the KNN method was: accuracy = 0.60, specificity = 0.66 and recall 0.51.

It can be seen, Table 1, that there was a statistically significant difference between, i.e. p value was less than 0.01, the classifier for all the metrics (accuracy, specificity and recall).

A post hoc test, Tukey multiple comparisons of means, was performed to verify the statistical difference between the results and it is shown on Table 2.
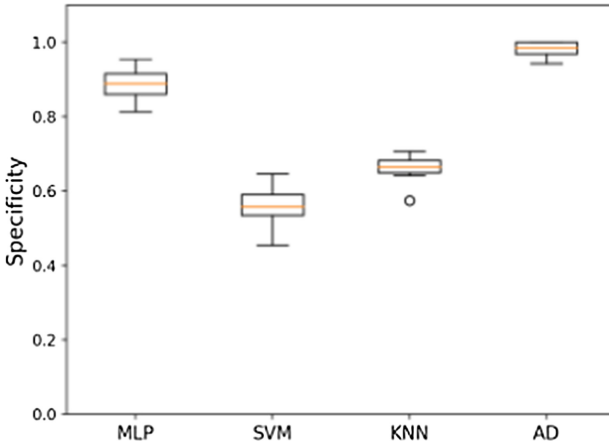
Fig. 5. Box plot of the performance for different classification models and the metric is Specificity. The models are MLP; SVM with kernel RBF; KNN; and AD. Image Source: Own Authors
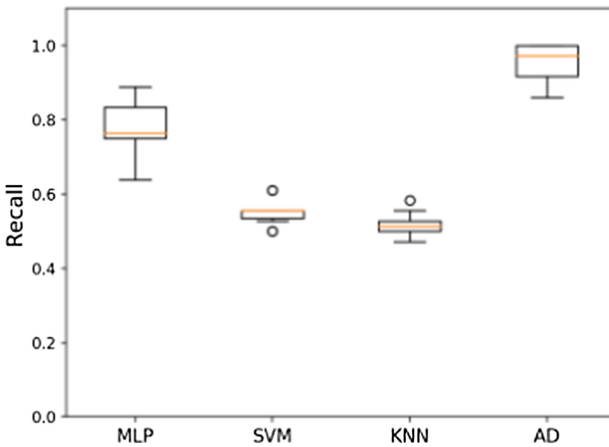

Fig. 6. Box plot of the performance for different classification models and the metric is Recall. The models are MLP; SVM with kernel RBF; KNN; and AD. Image Source: Own Authors

It can be seen that for almost all combinations the difference was statistically relevant (p-value < 0.01), i.e. the decision tree was the best estimate model whereas MLP was the second better. KNN was the third better algorithm when using the accuracy and specificity metrics. However, there was no statistically significant difference between KNN and SVM for the recall metric.

Table 1. ANOVA for Classification Models. Df stands for Degrees of Freedom, Df.res is the residue of the Degrees of Freedom, Sum Sq is the sum of squares, Sum Sq.res is the residue of the sum of squares, Mean Sq are the Mean Squares and Pr(>F) is the p-value.

Metric	Df	Df.res	Sum Sq	Sum Sq.res	Mean Sq	Pr(>F)
Accuracy	3	36	1,247	0,054	0,415	$2 * 10^{-16}$
Recall	3	36	1,275	0,093	0,425	$2 * 10^{-16}$
Specificity	3	36	1,147	0,057	0,382	$2 * 10^{-16}$

Table 2. Post hoc test - Tukey multiple comparisons of means

Algorithms	P (Accuracy)	P (Specificity)	P (Recall)
KNN-AD	0,0000000	0,0000000	0,0000000
MLP-AD	0,0000001	0,0000366	0,0000000
SVM-AD	0,0000000	0,0000000	0,0000000
MLP-KNN	0,0000000	0,0000000	0,0000000
SVM-KNN	0,0018110	0,0000126	0,4004624
SVM-MLP	0,0000000	0,0000000	0,0000000

5 Conclusion

This work introduced a software to measure the performance of subjects on a tracing task. This task and setup replicated an interesting previous work from motor skill learning. However, the present work went beyond the replicated study. The collected data with the software was also used on classification algorithms rather than only to indicate motor learning during the task, which was also demonstrated. This data was employed to predict the block a subject is using his/her dataset.

The results suggest that is possible to apply classification models to correlate the movement time and the percentage error with the number of blocks. This approach, which is proposed on this work, might be useful to estimate the number of trials, blocks, and sessions a patient requires to rehabilitate. Future works will be devoted to have more subjects as well as to use the proposed approach with patient.

References

1. Abe, M., Schambra, H., Wassermann, E.M., Luckenbaugh, D., Schweighofer, N., Cohen, L.G.: Reward improves long-term retention of a motor memory through induction of offline memory gains. Curr. Biol. **21**(7), 557–562 (2011)

2. Espírito-Santo, H., Pires, C.F., Garcia, I.Q., Daniel, F., Silva, A.G.d., Fazio, R.L.: Preliminary validation of the Portuguese Edinburgh Handedness Inventory in an adult sample. Appl. Neuropsychol. Adult **24**(3), 275–287 (2017)
3. Forkert, N.D., Verleger, T., Cheng, B., Thomalla, G., Hilgetag, C.C., Fiehler, J.: Multiclass support vector machine-based lesion mapping predicts functional outcome in ischemic stroke patients. PloS ONE **10**(6), e0129569 (2015)
4. Gold, J.I., Shadlen, M.N.: The neural basis of decision making. Annu. Rev. Neurosci. **30**, 235–274 (2007)
5. Jordan, M.I., Mitchell, T.M.: Machine learning: trends, perspectives and prospects. Science **349**(6245), 255–260 (2015)
6. Julianjatsono, R., Ferdiana, R., Hartanto, R.: High-resolution automated Fugl-Meyer Assessment using sensor data and regression model. In: 3rd International Conference on Science and Technology-Computer (ICST), pp. 28–32. IEEE (2017)
7. Kitago, T., Krakauer, J.W.: Motor learning principles for neurorehabilitation. Handb. Clin. Neurol. **110**, 93–103 (2013)
8. Lin, W.Y., et al.: Predicting post-stroke activities of daily living through a machine learning-based approach on initiating rehabilitation. Int. J. Med. Inf. **111**, 159–164 (2018)
9. Marsland, S.: Machine Learning: An Algorithmic Perspective. CRC Press, Boca Raton (2015)
10. Norvig, P., Russell, S.: Inteligência Artificial, vol. 1, 3rd edn. Elsevier, Rio de Janeiro (2014)
11. Pedregosa, F., et al.: Scikit-learn: machine learning in Python. J. Mach. Learn. Res. **12**, 2825–2830 (2011)
12. Prichard, G., Weiller, C., Fritsch, B., Reis, J.: Effects of different electrical brain stimulation protocols on subcomponents of motor skill learning. Brain Stimul. **7**(4), 532–540 (2014)
13. Reis, J., et al.: Noninvasive cortical stimulation enhances motor skill acquisition over multiple days through an effect on consolidation. Proc. Nat. Acad. Sci. **106**(5), 1590–1595 (2009)
14. Rossum, G.: Python reference manual. Technical report, Amsterdam, The Netherlands (1995)
15. Sailer, U., Flanagan, J.R., Johansson, R.S.: Eye-hand coordination during learning of a novel visuomotor task. J. Neurosci. **25**(39), 8833–8842 (2005)
16. Sale, P., et al.: Predicting motor and cognitive improvement through machine learning algorithm in human subject that underwent a rehabilitation treatment in the early stage of stroke. J. Stroke Cerebrovasc. Dis. **27**(11), 2962–2972 (2018)
17. Sato, A., Fujita, T., Ohashi, Y., Yamamoto, Y., Suzuki, K., Otsuki, K.: A prediction model for activities of daily living for stroke patients in a convalescent rehabilitation ward. J. Allied Health Sci. **7**(1), 1–6 (2016)
18. SDH/PR: Cartilha do censo 2010 pessoas com deficiência. Secretaria de Direitos Humanos da Presidência da República (2012). http://www.pessoacomdeficiencia.gov.br/app/sites/default/files/publicacoes/cartilha-censo-2010-pessoas-com-deficienciareduzido.pdf
19. Shinners, P.: Pygame (2011). http://pygame.org/
20. Sonoda, S., Saitoh, E., Nagai, S., Okuyama, Y., Suzuki, T., Suzuki, M.: Stroke outcome prediction using reciprocal number of initial activities of daily living status. J. Stroke Cerebrovasc. Dis. **14**(1), 8–11 (2005)
21. Sprint, G., Cook, D.J., Weeks, D.L., Borisov, V.: Predicting functional independence measure scores during rehabilitation with wearable inertial sensors. IEEE Access **3**, 1350–1366 (2015)

22. Triesch, J., Ballard, D.H., Hayhoe, M.M., Sullivan, B.T.: What you see is what you need. J. Vis. **3**(1), 9 (2003)
23. Tsuji, T., Liu, M., Sonoda, S., Domen, K., Chino, N.: The stroke impairment assessment set: its internal consistency and predictive validity. Arch. Phys. Med. Rehabil. **81**(7), 863–868 (2000)
24. Wolpert, D.M., Diedrichsen, J., Flanagan, J.R.: Principles of sensorimotor learning. Nat. Rev. Neurosci. **12**(12), 739 (2011)