Multiple Linear Regression in Predicting Motor Assessment Scale of Stroke Patients

Sulaiman Mazlan, Hisyam Abdul Rahman, Babul Salam Ksm Kader, Che Fai Yeong, and Nurul Aisyah Mohd Rostam Alhusni

Final Published Version deposited by Coventry University’s Repository

Original citation & hyperlink:
Mazlan, S., Rahman, H.A., Kader, B.S.K., Yeong, C.F. and Alhusni, N.A.M.R., 2021. Multiple Linear Regression in Predicting Motor Assessment Scale of Stroke Patients. *International Journal of Integrated Engineering*, 13(6), pp.330-338. [https://dx.doi.org/10.30880/ijie.2021.13.06.029](https://dx.doi.org/10.30880/ijie.2021.13.06.029)

DOI 10.30880/ijie.2021.13.06.029

ISSN 2229-838X
ESSN 2600-7916

Publisher: Penerbit UTHM

This work is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License](https://creativecommons.org/licenses/by-nc-sa/4.0/).
Multiple Linear Regression in Predicting Motor Assessment Scale of Stroke Patients

Sulaiman Mazlan¹, Hisyam Abdul Rahman¹*, Babul Salam Ksm Kader Ibrahim², Yeong Che Fai³, Nurul Aisyah Mohd Rostam Alhusni⁴

¹Advanced Mechatronics Research Group (ADMIRE), Universiti Tun Hussein Onn Malaysia, Batu Pahat, 86400, MALAYSIA
²Faculty of Engineering, Environment and Computing, Coventry University, Coventry, CV1 5FB, UNITED KINGDOM
³Center of Artificial Intelligence and Robotics, Universiti Teknologi Malaysia, Skudai, 81310, MALAYSIA
⁴Occupational Therapy Department, SOCSO Tun Razak Rehabilitation Centre, Melaka, 75450, MALAYSIA

*Corresponding Author
DOI: https://doi.org/10.30880/ijie.2021.13.06.029
Received 09 March 2021; Accepted 29 June 2021; Available online 31 August 2021

Abstract: The Multiple Linear Regression (MLR) is a predictive model that was commonly used to predict the clinical score of stroke patients. However, the performance of the predictive model slightly depends on the method of feature selection on the data as input predictor to the model. Therefore, appropriate feature selection method needs to be investigated in order to give an optimum performance of the prediction. This paper aims (i) to develop predictive model for Motor Assessment Scale (MAS) prediction of stroke patients, (ii) to establish relationship between kinematic variables and MAS score using a predictive model, (iii) to evaluate the prediction performance of a predictive model based on root mean squared error (RMSE) and coefficient of determination R². Three types of feature selection methods involve in this study which are the combination of all kinematic variables, the combination of the best four or less kinematic variables, and the combination of kinematic variables based on p < 0.05. The prediction performance of MLR model between two assessment devices (iRest and ReHAD) has been compared. As the result, MLR model for ReHAD with the combination of kinematic variables that has p < 0.05 as input predictor has the best performance with Draw I (RMSEte = 1.9228, R² = 0.8623), Draw Diamond (RMSEte = 2.6136, R² = 0.7477), and Draw Circle (RMSEte = 2.1756, R² = 0.8268). These finding suggest that the relationship between kinematic variables and MAS score of stroke patients is strong, and the MLR model with feature selection of kinematic variables that has p < 0.05 is able to predict the MAS score of stroke patients using the kinematic variables extracted from the assessment device.

Keywords: Multiple linear regression, robotic, rehabilitation, upper limb, stroke

1. Introduction

Upper limb motor dysfunction is one of the most relevant functions impaired by stroke, can lead to limitations of function and dramatically reduce the quality of life of stroke patients [1–4]. Due to the motor dysfunction, upper limb disability has subsequent effects on independence in daily activities, destination for discharge, return to work, quality of
life and mood [5–7]. It is important that stroke patient to undergo the upper limb rehabilitation to recover from upper limb disability. The intention of the upper limb rehabilitation is to improve the functional use of the arm in order to enable the person to carry out productive activities in real life. Improved motor function also contributes to the patients’ satisfaction, independence and improve quality of life [8].

Various types of clinical scales such as Fugl-Mayer (FMA) [9,10], Manual Muscle Test (MMT) [11,12], or Motor Assessment Scale (MAS) [13,14] are commonly used by physiotherapist to evaluate motor function of stroke patients during the rehabilitation program. However, the evaluation of the motor function using conventional clinical scales is challenging due to the time and limitation of resources [15]. In addition, the scoring systems are often subjective, lack reliability and heavily dependent on the ability of the skilled physiotherapist to provide only rough motor function estimates [13,16]. Nowadays, various types of upper limb assessment device for stroke rehabilitation have been developed to assist physiotherapists during rehabilitation program [17–23]. These upper limb assessment devices provide precise measurement of patient’s motor sensory performance which can have a beneficial impact on the rehabilitation outcome [15,24]. Kinematic variables evaluated by the assessment device have been used as independent variables in multivariate analysis for predicting the patient’s clinical score [15].

The extraction the relevant part of information for a large dataset to predict the clinical scale of stroke patients can be performed with different types of multivariate analysis methods. The Multiple Linear Regression (MLR) approach is commonly used method to obtain a linear input output model for a given dataset [25]. However, the performance of the predictive model slightly depends on the feature selection method used. Therefore, appropriate feature selection method needs to be investigated in order to give an optimum performance of the prediction. The main objective of this study is to compare the prediction performance between two assessment devices (iRest and ReHAD) using MLR analysis. This paper aims (i) to develop predictive model for MAS score prediction of stroke patients, (ii) to establish relationship between kinematic variables and MAS score using a predictive model, (iii) to evaluate the predictive accuracy of a predictive model based on root mean squared error (RMSE) and coefficient of determination R².

2. Research Method

The data collection is conducted following the ethical approval granted by the Universiti Tun Hussein Onn Malaysia (UTHM) Research Ethics Committee. Subjects have been selected by the occupational therapists in SOCSO Tun Razak Rehabilitation Centre, based on the inclusion criteria of the study which the upper limb stroke patients with a MAS score of 3 and above. All subjects received conventional physiotherapy daily. Each subject’s motor sensory function was evaluated at the end of the study using the MAS. Subjects participated in a 30-minutes robotic assessment, including 10 minutes for each assessment module. The robotic assessment start with Draw I, Draw Diamond and Draw Circle module in sequences, where the set-up of the experiment was the same as the previous study [16,26]. The grasping system for iRest was used to measure hand opening and closing movement while the grasping system for ReHAD was used to measure the hand grip force. Subjects were asked to grasp the handle of the assessment device and their affected hands were covered by Velcro band.

General idea of the research methodology shows in Fig. 1. The raw data from the assessment device will be processed through feature extraction stage. Twelve kinematic variables will be produced as the output of the feature extraction stage. After that, MLR multivariate calibration will be used for modelling the data and generate prediction of MAS score for each stroke patients.

![Fig. 1 - The flows of MAS score prediction](image)

2.1 Data Collection

The raw data were extracted from the developed assessment devices (iRest and ReHAD). The number of 50 stroke patients (36 male and 14 female) that has upper limb disability participated to perform robotic assessment process in this study. The data for each patient has been extracted from the assessment device including time, position, and grip force. All stroke patients required to perform the assessment task for three trials. Total of 150 data set had been produced after the assessment process for 50 stroke patients.

2.2 Feature Extraction

Feature extraction is the process of reducing the dimensions of the raw data collected with the assessment device without compromising the data information that has been collected. Raw data taken from the assessment device will be processed and evaluated as kinematic variables. A systematic review shows there are various types of kinematic variables that have been used as indicators to assess patient motor performance [24]. In this study, the kinematic variables have

331
been calculated using MATLAB software. Twelve kinematic variables were extracted from the rehabilitation device: Total movement time, reaction time, stability time, mean velocity, time to peak velocity, peak velocity, path ratio, hit-wall score, number of peaks speed, trajectory error, target reached, and grasping. All the calculation for the kinematic variables were referred from the previous study [16].

2.3 Feature Selection

Feature selection is the process of selecting the combination of predictor variables that most contributes to the forecast model. This study uses three types of feature selection methods in order to observe the best input combination to the MLR model. The first method is to use all kinematic variables as the input predictor.

Several study shows that four kinematic variables were high enough to result a good predictive performance of a regression model [16,27]. In addition, a study conducted using MIT-Manus used twenty kinematic variables as an independent variable in MLR model, but only four kinematic variables were retained and resulting the best performance of prediction [28]. Therefore, selection of the best four or less combination of the kinematic variables has been selected as the second feature selection method in order to evaluate the performance of the linear regression model for predicting the clinical scores.

A study used univariate regression to identify the kinematic variables with p-value lower than 0.2 for the multiple regression model [29]. However, the study only retained the kinematic value with p < 0.05 for the final models as it has more significant contribution to the regression model. Therefore, selection of the kinematic variables with p < 0.05 has been selected as the third feature selection method in order to evaluate the performance of the linear regression model for predicting the clinical scores.

2.4 Multiple Linear Regression

Multiple Linear Regression (MLR) approaches is the basic and simple method for experimental and data processing in analytical data [30]. MLR is a powerful statistical tool finding relationships between one dependent and multiple independent variables [31–33]. In MLR, the dependent variables y is linearly correlated to multiple independent variables $x_1, x_2, ..., x_n$. The multiple linear regression model as in Eq. (1) as follow:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_n x_n$$

where, y is dependent variable, x is independent variables, $\beta_0$ is bias, $\beta_i$, $\beta_2$, ..., $\beta_n$ are the coefficient of independent variables. These parameters are estimated by training the samples. Most analysis to predict the upper limb assessment in stroke rehabilitation using MLR shows strong correlation with the clinical scales [13,27,32].

2.5 Validation

Each stroke patients required to perform three trials for each assessment module. Two of the trials will be used as the training data set while the other trial will be holded out as the unseen validation data set for the MLR validation. The root mean square error of training (RMSE$_{tr}$), root mean square error of testing (RMSE$_{te}$) and coefficient of determination of prediction has been used to represent prediction accuracy capacity of developed model. The RMSE$_{tr}$ was calculated in Eq. (2) as follow:

$$RMSE_{tr} = \sqrt{\frac{\sum_{i=1}^{n} (\hat{y}_{tr} - y_{tr})^2}{n}}$$

Where $\hat{y}_{tr}$ represent the predicted assessment score from training data set, $y_{tr}$ denote the reference clinical score from training data set, n represent the total number of training samples. The root mean squared error of testing (RMSE$_{te}$) was used to measures the accuracy of the predictions of the predictive model with new unseen of data set can be computed in Eq. (3) as follow:

$$RMSE_{te} = \sqrt{\frac{\sum_{i=1}^{n} (\hat{y}_{te} - y_{te})^2}{n}}$$

Where $\hat{y}_{te}$ represent the predicted assessment score from testing data set, $y_{te}$ denote the reference clinical score from testing data set, n represent the total number of testing samples. The coefficient of determination of prediction used was interpreted as the proportion of variance in the prediction of the reference value of regression analysis is defined as in Eq. (4).
Where $y_\overline{}$ represent the mean of reference data, $\hat{y}_i$ denote the predicted assessment score and $y_i$ denote the reference clinical score. The $R^2$ of the predictive models were measured to describe the relationship between robotic assessment score and clinical assessment score.

3. Results and Discussion

The results were discussed in this section including the selection of kinematic variables combination as the predictor for MLR predictive model based on feature selection method. The performance result of all MLR predictive model for all assessment modules were discussed at the end of this section.

3.1 Feature Selection

This sub-section shows the result of feature selection method including the combination of four or less kinematic variables, and the combination of kinematic variables selected based on $p$-value below 0.05.

3.1.1 Combination of Four or Less Kinematic Variables

The best combination of kinematic variables was determined using leave one out cross validation (LOOCV) approach. One data point was released in turn, the remaining data were used to fit the predictive model. The error between the predicted value of the unused data point and the actual value was calculated as root mean square error of LOOCV ($\text{RMSE}_{cv}$). The combination that produced the least $\text{RMSE}_{cv}$ value was selected using an exhaustive search of all possible combinations. The $\text{RMSE}_{cv}$ for the predictive model derived from LOOCV process tabulated in Table 1 and Table 2 for iRest and ReHAD respectively. The results show the combination of the best four or less kinematic variables in predicting the MAS score based on the $\text{RMSE}_{cv}$ value.

### Table 1 - $\text{RMSE}_{cv}$ values from different combination of kinematic variables for iRest

| Comb. of variables | Draw I Kinematic variables | RMSE$_{cv}$ | Draw Diamond Kinematic variables | RMSE$_{cv}$ | Draw Circle Kinematic variables | RMSE$_{cv}$ |
|--------------------|---------------------------|-------------|----------------------------------|-------------|---------------------------------|-------------|
| 4                  | Movement time, Path ratio, Grasping, Target Reached | 2.6986      | Movement time, Reaction time, Grasping, Target Reached | 2.8920      | Reaction time, Hit wall score, Grasping, Target Reached | 2.7821      |
| 3                  | Movement time, Grasping, Target Reached | 2.7141      | Movement time, Grasping, Target Reached | 2.9209      | Hit wall score, Grasping, Target Reached | 2.8108      |
| 2                  | Grasping, Target Reached | 2.8131      | Grasping, Target Reached | 3.0068      | Grasping, Target Reached | 2.8613      |
| 1                  | Grasping | 3.3311      | Grasping | 3.2365      | Grasping | 3.0283      |

### Table 2 - $\text{RMSE}_{cv}$ values from different combination of kinematic variables for ReHAD

| Comb. of variables | Draw I Kinematic variables | RMSE$_{cv}$ | Draw Diamond Kinematic variables | RMSE$_{cv}$ | Draw Circle Kinematic variables | RMSE$_{cv}$ |
|--------------------|---------------------------|-------------|----------------------------------|-------------|---------------------------------|-------------|
| 4                  | Peak velocity, Hit wall score, Grasping, Target Reached | 2.0012      | Stability time, Hit wall score, Grasping, Trajectory error | 2.4512      | Movement time, Stability time, Hit wall score, Grasping | 2.2718      |
| 3                  | Mean velocity, Hit wall score, Grasping | 2.1485      | Stability time, Grasping, Trajectory error | 2.5441      | Stability time, Peak velocity, Grasping | 2.3540      |
| 2                  | Hit wall score, Grasping | 2.2963      | Stability time, Peak velocity | 2.6544      | Stability time | 2.6018      |
| 1                  | Stability time | 2.6744      | Stability time | 2.7577      | Stability time | 2.8069      |

Based on iRest result in Table 1. The minimum $\text{RMSE}_{cv}$ value was 2.6986 for Draw I model with a combination of four kinematic variables (Movement time, Path ratio, Grasping, and Target Reached). 2.8920 for Draw Diamond model with a combination of four kinematic variables (Movement time, Reaction time, Grasping, and Target Reached), and
Eight kinematic variables (Movement time, Stability time, Mean velocity, Peak spring, and Trajectory error), the kinematic variables exceed the specified kinematic variables for ReHAD is higher compared to the iRest in each assessment modules due to most of the kinematic variables were selected for Draw I module except for two kinematic variables (Movement time and Trajectory error). Meanwhile, only Trajectory error was excluded from the combination of kinematic variables for Draw Diamond module, two kinematic variables (Reaction time and Trajectory error) were excluded from the combination of kinematic variables for Draw Circle module. Trajectory error was the only single kinematic variables that were ignored from the kinematic variable’s combination for ReHAD assessment modules. Based on Table 3 and Table 4, the number of selected kinematic variables for ReHAD is higher compared to the iRest in each assessment modules due to most of the kinematic variables exceed the specified inclusion criteria ($p < 0.05$).

### 3.1.2 Combination of Kinematic Variables ($p < 0.05$)

Pearson's Linear Correlation Coefficient was used to determine the correlation between pairs of all independent variables and dependent variables. The kinematic variables below 0.05 were selected as the input combination of variables for generating the prediction model. Table 3 shows the $p$-value of each kinematic variables for three assessment modules.

**Table 3 - $p$-value of each kinematic variable for iRest**

| Kinematic variables | Draw I    | Draw Diamond | Draw Circle |
|---------------------|-----------|--------------|-------------|
| Movement time       | 0.0052    | 0.6899       | 0.8108      |
| Stability time      | 9.9921e-17 | 3.2560e-21   | 4.0749e-27  |
| Reaction time       | 0.8907    | 6.5480e-06   | 2.1203e-04  |
| Mean velocity       | 1.3625e-11 | 5.5884e-05   | 2.5284e-04  |
| Peak velocity       | 0.0026    | 0.0171       | 0.2223      |
| Time to peak velocity | 0.9393 | 0.0037       | 2.0323e-05  |
| Hit wall score      | 0.6261    | 0.0132       | 0.3965      |
| Path ratio          | 1.4117e-07 | 4.8212e-08   | 7.8928e-07  |
| Smoothness          | 0.1035    | 0.0151       | 0.0497      |
| Grasping            | 1.6467e-35 | 1.9265e-37   | 7.6301e-42  |
| Trajectory error    | 0.0332    | 0.2467       | 0.4547      |
| Target Reached      | 7.7094e-25 | 2.8849e-15   | 4.7848e-16  |

### Table 4 - $p$-value of each kinematic variable for ReHAD

| Kinematic variables | Draw I    | Draw Diamond | Draw Circle |
|---------------------|-----------|--------------|-------------|
| Movement time       | 7.1707e-28 | 8.7223e-09   | 1.0076e-05  |
| Stability time      | 4.6872e-45 | 4.3427e-43   | 6.3080e-42  |
| Reaction time       | 0.0015    | 0.0013       | 0.0652      |
| Mean velocity       | 2.6448e-29 | 3.3948e-10   | 4.6993e-08  |
| Peak velocity       | 0.2649    | 1.1231e-05   | 4.1691e-08  |
| Time to peak velocity | 0.0197 | 0.0026       | 2.8253e-05  |
| Hit wall score      | 7.0459e-18 | 0.0186       | 0.0022      |
| Path ratio          | 1.0887e-06 | 1.4114e-04   | 2.3391e-05  |
| Smoothness          | 0.4554    | 2.6759e-07   | 5.8290e-11  |
| Grasping            | 1.6791e-36 | 1.3784e-29   | 1.1655e-35  |
| Trajectory error    | 0.0915    | 0.2016       | 0.1582      |
| Target Reached      | 1.1630e-11 | 0.0011       | 0.0015      |

Table 3 shows the result for iRest. Eight kinematic variables (Movement time, Stability time, Mean velocity, Peak velocity, Path ratio, Grasping, Trajectory error, and Target reached) were selected for Draw I module. All kinematic variables were selected for Draw Diamond module except for two kinematic variables (Movement time and Trajectory error). Besides, four kinematic variables (Movement time, Peak velocity, Hit wall score, and Trajectory error) were excluded from the combination of kinematic variables for Draw Circle module due to $p > 0.05$. Based on Table 4, all kinematic variables were selected for Draw I module except three kinematic variables (Peak velocity, Smoothness, and Trajectory error). Meanwhile, only Trajectory error was excluded from the combination of kinematic variables for Draw Diamond module, two kinematic variables (Reaction time and Trajectory error) were excluded from the combination of kinematic variables for Draw Circle module. Trajectory error was the only single kinematic variables that were ignored from the kinematic variable’s combination for ReHAD assessment modules. Based on Table 3 and Table 4, the number of selected kinematic variables for ReHAD is higher compared to the iRest in each assessment modules due to most of the kinematic variables exceed the specified inclusion criteria ($p < 0.05$).
3.2 Performance prediction of MLR model

The performance of the MLR predictive models were observed from the value of RMSE\textsubscript{te} and \(R^2\). The performance of the MLR predictive model shown in Table 5 and Table 6 for iRest and ReHAD respectively.

Table 5 - The performance of MLR model for iRest

| Features selection | Module  | Training | Testing |
|--------------------|---------|----------|---------|
| All kinematic variables | Draw I | 2.4846 | 0.8001 | 2.6232 | 0.7921 |
|                      | Draw D | 2.6755 | 0.7683 | 2.8553 | 0.7374 |
|                      | Draw C | 2.3388 | 0.8228 | 3.2943 | 0.6642 |
| Best 4 combination   | Draw I | 2.5763 | 0.7852 | 2.6379 | 0.7758 |
|                      | Draw D | 2.7717 | 0.7513 | 2.8868 | 0.7312 |
|                      | Draw C | 2.4422 | 0.8069 | 3.1015 | 0.6934 |
| \(p\_value < 0.05\)  | Draw I | 2.5253 | 0.7935 | 2.5952 | 0.7882 |
|                      | Draw D | 2.7072 | 0.7627 | 2.8945 | 0.7305 |
|                      | Draw C | 2.7325 | 0.7583 | 2.7544 | 0.7550 |

Table 6 - The performance of MLR model for ReHAD

| Features selection | Module  | Training | Testing |
|--------------------|---------|----------|---------|
| All kinematic variables | Draw I | 1.4948 | 0.9166 | 2.1968 | 0.8221 |
|                      | Draw D | 1.9117 | 0.8636 | 2.5538 | 0.7592 |
|                      | Draw C | 2.0198 | 0.8477 | 2.1606 | 0.8276 |
| Best 4 combination   | Draw I | 1.9273 | 0.8614 | 1.9591 | 0.8571 |
|                      | Draw D | 2.2883 | 0.8046 | 2.5173 | 0.7672 |
|                      | Draw C | 2.1082 | 0.8341 | 2.3366 | 0.7997 |
| \(p\_value < 0.05\)  | Draw I | 1.8975 | 0.8656 | 1.9228 | 0.8623 |
|                      | Draw D | 2.0603 | 0.8416 | 2.6136 | 0.7477 |
|                      | Draw C | 2.0404 | 0.8446 | 2.1756 | 0.8268 |

Based on Table 5, the results show that all feature selection method has RMSE\textsubscript{te} value below 3.3 for each assessment module using iRest. The first feature selection method where all kinematic variables involved in regression analysis, Draw I modules score the best prediction result (RMSE\textsubscript{te} = 2.6232, \(R^2 = 0.7921\)) compared to the other two modules. Draw C has improved the prediction performance (RMSE\textsubscript{te} = 3.1015, \(R^2 = 0.6934\)) in the second feature selection method where only the best four kinematic variables were selected. However, prediction performance of Draw I and Draw D modules were decreased compared to the first feature selection method. The third feature selection method where only the kinematic variable that has \(p\_value < 0.05\) were retained for the regression analysis, Draw I modules shows the best prediction performance (RMSE\textsubscript{te} = 2.5952, \(R^2 = 0.7882\)) followed by Draw C and Draw D modules. The performance of MLR model for the iRest shows that Draw I modules has the excellent performance for all feature selection method involved.

Based on Table 6, Draw C module scores the excellent prediction performance (RMSE\textsubscript{te} = 2.1968, \(R^2 = 0.8221\)) as compared to the other two modules for the first feature selection method where all kinematic variables involved in regression analysis. Besides, Draw D module has the worst prediction performance due to higher value of RMSE\textsubscript{te} and lower value of \(R^2\). The second feature selection method where only the best four kinematic variables were selected as the input for MLR model, Draw I module has the best prediction result (RMSE\textsubscript{te} = 1.9591, \(R^2 = 0.8571\)) followed by Draw C and Draw D modules. In addition, the performance of Draw I module has increased by 10.82% of RMSE\textsubscript{te} value and 4.26% of \(R^2\) value compared to the first feature selection method. The third feature selection method where only the kinematic variable that has \(p\_value < 0.05\) were retained for the regression analysis, Draw I modules shows the best prediction performance (RMSE\textsubscript{te} = 1.9228, \(R^2 = 0.8623\)) followed by Draw C and Draw D modules. The performance of MLR model for ReHAD shows that Draw I modules has the admirable performance with two out of three feature selection method involved in the MLR analysis.
Since the main objective of this paper is to compare the performance of MLR model for both assessment devices, the MLR model for ReHAD device resulted a better performance compared to the MLR model for iRest device. This is proven by the lower value of RMSE in and higher value of $R^2$. Furthermore, the third feature selection method where only the kinematic variable that has $p$-value below than 0.05 were retained for the regression analysis shows the magnificent performance compared to the other two feature selection methods. The results indicate that MLR model for ReHAD with third feature selection method has more robustness in testing new unseen test sets samples of kinematic variables. Fig. 2 shows the correlation between predicted MAS score and MAS score for ReHAD with feature selection of kinematic variables ($p < 0.05$) as the input predictor to MLR model. The predicted MAS score and MAS score values showed statistically significant ($p < 0.05$) correlations in all cases. In addition, the training and the validation model for all assessment modules showed that the predicted MAS score were positively correlated with the MAS score. However, MLR model is going to be ineffective for the system with nonlinear data due to limitation of MLR as linear predictive model. The performance of the prediction should be improved by including non-linear or hybrid predictive model in the future.

4. Conclusion

A study has been conducted using two assessment devices which are iRest and ReHAD in order to predict the clinical scale score using Multiple Linear Regression (MLR). To sum, MLR is promising to predict the motor assessment scale (MAS) score from the extracted kinematic variables of stroke patients. The results show MLR model for ReHAD has a better performance of prediction compared to iRest. In addition, optimization in feature selection method is crucial to improve the prediction performance. Finding also shows that feature selection of kinematic variables that has $p$-value below than 0.05 as input variables for the MLR model give excellent performance of prediction.

Acknowledgement

The authors would like to thank the physiotherapist from SOCSO Tun Razak Rehabilitation Centre for providing valuable feedback and recommendation in this study and also to Ministry of Higher Education Malaysia (MOHE) and Universiti Tun Hussein Onn Malaysia (UTHM) for their supports under FRGS-RACER Vot K144 and GPPS Vot H356 grants.

References

[1] Ferreira, F.M.R.M., Chaves, M.E.A., Oliveira, V.C., Van Petten, A.M.V.N., and Vimieiro, C.B.S. (2018). Effectiveness of robot therapy on body function and structure in people with limited upper limb function: A systematic review and meta-analysis. Public Library of Science. 13 (7), e0200330
[2] Ang, B.W.K. and Yeow, C. (2019). Design and Characterization of a 3D Printed Soft Robotic Wrist Sleeve with 2 DoF for Stroke Rehabilitation. in: 2019 2nd IEEE Int. Conf. Soft Robot., pp. 577–582
[3] Zimmermann, Y., Forino, A., Rieno, R., and Hutter, M. (2019). ANYexo: A Versatile and Dynamic Upper-Limb Rehabilitation Robot. IEEE Robotics and Automation Letters. 4 (4), 3649–3656
[4] Chen, Z., Wang, C., Fan, W., Gu, M., Yasin, G., Xiao, S., et al. (2020). Robot-Assisted Arm Training versus Therapist-Mediated Training after Stroke: A Systematic Review and Meta-Analysis. Journal of Healthcare Engineering. 2020 810867
[5] Signal, N.E.J., McLaren, R., Rashid, U., Vandal, A., King, M., Almesfer, F., et al. (2020). Haptic Nudges Increase Affected Upper Limb Movement During Inpatient Stroke Rehabilitation: Multiple-Period Randomized Crossover Study. JMIR Mhealth Uhealth. 8 (7), e17036
[6] Morris, J.H., van Wijck, F., Joice, S., and Donaghy, M. (2013). Predicting health related quality of life 6 months after stroke: the role of anxiety and upper limb dysfunction. Disability and Rehabilitation. 35 (4), 291–299

Fig. 2 - Predicted MAS score versus MAS score using ReHAD with feature selection of ($p < 0.05$) of the kinematic variables for (a) Draw I module, (b) Draw D module and (c) Draw C module.
Sulaiman et al., International Journal of Integrated Engineering Vol. 13 No. 6 (2021) p. 330-338

[7] Wang, Y.-C., Kapellusch, J., and Garg, A. (2014). Important factors influencing the return to work after stroke. Work. 47 553–559

[8] Barbara, B., Y., C.J., W., D.P., I., G.J., D., G.G., C., K.R., et al. (2005). Veterans Affairs/Department of Defense Clinical Practice Guideline for the Management of Adult Stroke Rehabilitation Care. Stroke. 36 (9), 2049–2056

[9] Colombo, R., Pisano, F., Delconte, C., Mazzone, A., Grioni, G., Castagna, M., et al. (2017). Comparison of exercise training effect with different robotic devices for upper limb rehabilitation: a retrospective study. Eur J Phys Rehabil Med. 53 (2), 240–248

[10] Rech, K.D., Salazar, A.P., Marchese, R.R., Schifino, G., Cimolin, V., and Pagnussat, A.S. (2020). Fugl-Meyer Assessment Scores Are Related With Kinematic Measures in People with Chronic Hemiparesis after Stroke. Journal of Stroke and Cerebrovascular Diseases. 29 (1),

[11] Ciesla, N., Dinglas, V., Fan, E., Kho, M., Kuramoto, J., and Needham, D. (2011). Manual Muscle Testing: A Method of Measuring Extremity Muscle Strength Applied to Critically Ill Patients. Journal of Visualized Experiments: JoVE, 50

[12] Bohannon, R.W. (2019). Considerations and Practical Options for Measuring Muscle Strength: A Narrative Review. BioMed Research International. 2019 8194537

[13] Rahman, H.A. (2016). Non-motorized Three Degree of Freedom Assessment Tool for Stroke Patients, Universiti Teknologi Malaysia, 2016

[14] Barker, R.N., Brauer, S.G., and Carson, R.G. (2008). Training of Reaching in Stroke Survivors With Severe and Chronic Upper Limb Paresis Using a Novel Nonrobotic Device. Stroke. 39 (6), 1800–1807

[15] Tran, V.D., Dario, P., and Mazzoleni, S. (2018). Kinematic measures for upper limb robot-assisted therapy following stroke and correlations with clinical outcome measures: A review. Med Eng Phys. 53 13–31

[16] Abdul Rahman, H., Khor, K.X., Yeong, C.F., Su, E.L.M., and Narayanan, A.L.T. (2017). The potential of iRest in measuring the hand function performance of stroke patients. Bio-Medical Materials and Engineering. 28 (2), 105–116

[17] Ferreira, F.M.R.M., Rúbio, G. de P., Brandão, F.H. de L., Mata, A.M. da, Avellar, N.B.C. de, Bonfim, J.P.F., et al. (2020). Robotic Orthosis for Upper Limb Rehabilitation. Proceedings. 64 (1), 10

[18] Zhang, C., Li–Tsang, C.W., and Au, R.K. (2017). Robotic approaches for the rehabilitation of upper limb recovery after stroke: a systematic review and meta-analysis. International Journal of Rehabilitation Research. 40 (1), 19–28

[19] Mehrholz, J., Pohl, M., Platz, T., Kugler, J., and Elsner, B. (2015). Electromechanical and robot-assisted arm training for improving activities of daily living, arm function, and arm muscle strength after stroke. Cochrane Database of Systematic Reviews. (11), Cd006876

[20] Veerbeek, J.M., Langbroek-Amersfoort, A.C., van Wegen, E.E., Meskers, C.G., and Kwakkel, G. (2017). Effects of Robot-Assisted Therapy for the Upper Limb After Stroke. Neurorehabilitation and Neural Repair. 31 (2), 107–121

[21] Bertani, R., Melegari, C., De Cola, M.C., Bramanti, A., Bramanti, P., and Calabro, R.S. (2017). Effects of robot-assisted upper limb rehabilitation in stroke patients: a systematic review with meta-analysis. Neurological Sciences. 38 (9), 1561–1569

[22] Mazlan, S., Abdul Rahman, H., and Hanafi, D. (2020). A Review of Upper Limb Rehabilitation Robot. Journal of Tomography System and Sensor Application. 2 (1)

[23] Rahman, H.A. (2017). A Simple Upper Limb Rehabilitation Trainer. International Journal of Integrated Engineering. 9 (3), 39–43

[24] Mazlan, S., Abdul Rahman, H., Fai, Y., Ibrahim, B., and Huq, M. (2020). Kinematic variables for upper limb rehabilitation robot and correlations with clinical scales: A review. Bulletin of Electrical Engineering and Informatics. 9 (1), 75–82

[25] Muhamad Safiih, L., Ramlee, M., Gunalan, S., Zainuddin, N., Zakariya, R., Idris, M., et al. (2016). Improved the Prediction of Multiple Linear Regression Model Performance Using the Hybrid Approach: A Case Study of Chlorophyll-a at the Offshore Kuala Terengganu, Terengganu. Open Journal of Statistics. 06 789–804

[26] Rahman, H.A., Narayanan, A.L.T., Xiang, K.K., Ming, E.S.L., Fai, Y.C., and Khan, Q.I. (2015). iRest: Interactive rehabilitation and assessment tool. 2015 10th Asian Control Conference (ASCC). 1–6

[27] Zariffa, J., Kapadia, N., Kramer, J.L.K., Taylor, P., Alizadeh-Meghrazi, M., Zivanovic, V., et al. (2012). Relationship Between Clinical Assessments of Function and Measurements From an Upper-Limb Robotic Rehabilitation Device in Cervical Spinal Cord Injury. IEEE Transactions on Neural Systems and Rehabilitation Engineering. 20 (3), 341–350

[28] Bosecker, C., Dipietro, L., Volpe, B., and Igo Krebs, H. (2009). Kinematic Robot-Based Evaluation Scales and Clinical Counterparts to Measure Upper Limb Motor Performance in Patients With Chronic Stroke. Neurorehabilitation and Neural Repair. 24 (1), 62–69

[29] Hussain, N., Sunnerhagen, K.S., and Alt Murphy, M. (2019). End-point kinematics using virtual reality explaining upper limb impairment and activity capacity in stroke. Journal of NeuroEngineering and Rehabilitation. 16 (1), 82
[30] Balabin, R.M., Safieva, R.Z., and Lomakina, E.I. (2007). Comparison of linear and nonlinear calibration models based on near infrared (NIR) spectroscopy data for gasoline properties prediction. Chemometrics and Intelligent Laboratory Systems. 88 (2), 183–188

[31] Seber, G.A.F. and Lee, A.J. (2003). Linear Regression Analysis, Second Edition.

[32] Miler-Jerković, V., Djurić-Jovičić, M., Perović-Belić, M., Ječmenica-Lukić, M., Petrović, I.N., Radovanović, S.M., et al. (2014). Multiple regression analysis of repetitive finger tapping parameters. 2014 22nd Telecommunications Forum Telfor (TELFOR). 537–540

[33] Darmawan, M.F., Jamahir, N.I., Saedudin, R.D.R., and Kasim, S. (2018). Comparison between ANN and Multiple Linear Regression Models for Prediction of Warranty Cost. International Journal of Integrated Engineering. 10 (6), 193–196