2024 I-RIM Conference
October 25-27, Rome, Italy
ISBN: 9788894580556
10.5281/zenodo.14731084

Investigating the role of a mid-term clinical assessment in the prediction of robotic rehabilitation outcome in post-stroke survivors

Anastasios Tzepkenlis[1] Cristian Camardella[1] Antonio Frisoli [1]
Istituto di Intelligenza Meccanica, Scuola Superiore Sant'Anna, Pisa, Italy
{anastasios.tzepkenlis, cristian.camardella, antonio.frisoli}@santannapisa.it

***Abstract*—This research explores the impact of incorporating mid-term clinical assessments on predicting robotic rehabilitation outcomes for post-stroke patients using Random Forest models. While predicting motor recovery is complex and critical for personalizing therapies and reducing healthcare costs, adding mid-term assessments yielded limited improvements in prediction accuracy. The study compared predictions using only initial assessments versus both initial and mid-term assessments for three clinical scales: FMA, ARAT, and MI. Results showed that only ARAT predictions benefited from mid-term data, suggesting that further analysis is needed to fully understand the potential of mid-term assessments in enhancing predictive models.**

***Index Terms*—robotic rehabilitation, prediction, decision support, stroke, machine learning, mid-term assessment**

I. Introduction

The importance of predicting clinical outcomes before the end of a treatment has always been under the spotlight in the research and industrial product development phase. In particular it would allow for a huge cost reduction for the public health system and for a high personalization of therapies [1], thanks to a more conscious and fast reaction to different outcomes than what expected by the clinical personnel. Of course, predicting a clinical outcome is a complex task that involves many different factors and it requires a huge amount of data to be effective. One example of a potentially high impactful research in the field of predictive system concerns the stroke rehabilitation, which regards one third of the world population, as stroke is ranked second in worldwide mortality [2]. Post-stroke patients rehabilitation, although widely studied for years, still hides a high number of challenges. For instance, prediction of the rehabilitation outcome for post-stroke patients remains a complex and unresolved task, although many attempts have been made to build effective and useful estimations of the motor recovery process. In previous studies clinical scales scores of post-stroke patients at discharge have been predicted at the start of the therapy using pre-assessment scores, images, biomechanical data, demographics, and medical history [3]–[5]. In [6], [7], data coming from a rehabilitation robot have been used to extract kinematic and kinesiologic features to be inputs, together with clinical and demographics features [8], for a Random Forest model, showing good prediction capabilities. When planning a clinical trial protocol, a typical scenario consists in including two clinical assessments: a pre-assessment and a post-assessment. This is trivially linked to the need to understand the efficacy of the rehabilitation process. In rare cases, a mid-term assessment is included, although it could generate new insights and new data to improve current predictive models. The hypothesis is that, having a higher resolution of the evolution of clinical scale scores, it would provide machine learning models with new knowledge for a more accurate inference. In this study we investigated the usefulness of mid-term assessments (a clinical assessment performed halfway during the therapy) in enriching the input set, leading in some cases to more accurate prediction of the rehabilitation outcome.

Funding Agency: Horizon Europe Research & Innovation Programme

II. Materials and Methods

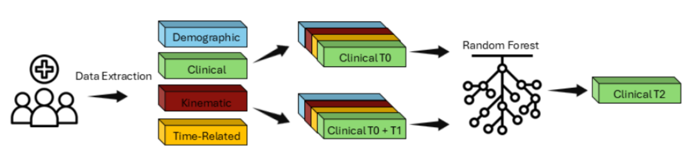
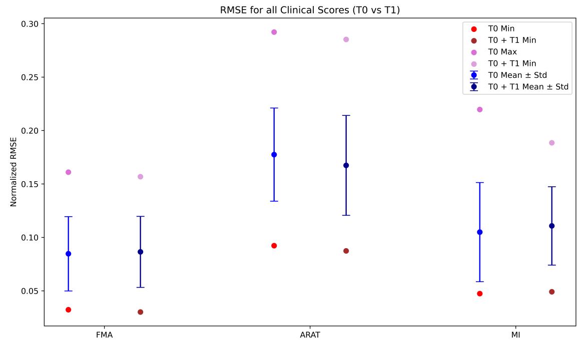
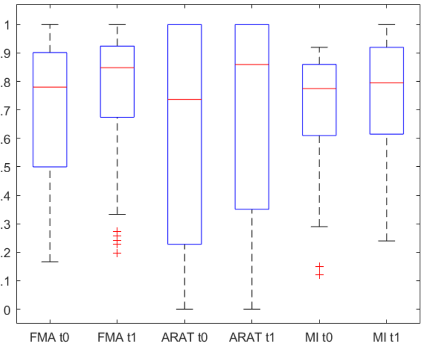


Fig. 1: Architecture of the model, together with the performed methods: two input set are used for the same model (with and without the mid-term assessment) and clinical scale scores at discharge (T2) are produced.

Demographic, clinical, and kinematic data were used as part of a previous dataset (registered at clinicaltrial.gov, code NCT05238389) collected during a clinical trial involving 40 post-stroke patients, as 4 out of the 44 original patients did not finish the rehabilitation therapy (age 67±12, FMA 46±16, ARAT 35±22, and MI 70±21 at T0), undergone to robotics rehabilitation with MOTORE robot and various serious games [9]. Kinematic data coming from each session were averaged intra-day, and features were extracted [7] from position, speed, and force measurements. From the rehabilitation progression and the amount of serious games played, the time-related features are generated (e.g. days since the beginning of the therapy, number of days between two rehabilitation sessions,



(a) Results comparing the prediction accuracy of the same Random Forest model, having the mid-term assessment in (left) and out(right) of the input set. The prediction accuracy is depicted for the FMA, ARAT, and MI clinical scales. Values on the y axes corresponds to the RMSE of the prediction, divided by the maximum of each scale.



(b) Comparison of clinical scale scores (FMA, ARAT, and MI) between start of the therapy (T0) and mid-term assessment (T1). Each boxplot depicts the distribution of scores among all the patients.

number of serious games completed in a day, etc.). Two conditions were investigated: in the first one we used only the clinical information from the initial clinical assessment (T0), while in the second one we add the clinical information from the intermediate clinical assessment (T1) after 20 days since the start of the rehabilitation (corresponding to half of therapy).

On both conditions we implemented Random Forest models [10], which are providing predictions for the outcome of the rehabilitation therapy on three different clinical scales, namely, FMA, ARAT and MI (Figure 1). On both conditions the models have been trained with the dataset being split into training (70%) and testing (30%), and a total of 30 different splits and 5 cross validations for both models have been conducted. All results in the following section are referring to the test set only. Considering the parameters of both Random Forest models (with the intermediate timepoint and without), we optimized the parameters into 300 estimators, 5 max depth, 20 minimum splits of each sample and 20 minimum samples per leaf. The normalized root-mean-square error (RMSE) function was introduced as the evaluation metric, allowing all three clinical scores to be presented on the same scale.

III. Results and Discussions

Results shown in Figure 2a compares the accuracy of the prediction of the proposed model with two different input sets: one with the mid-term assessment (MTA) as an additional clinical feature, and one without it. As it can be noted in the left plot, the mid-term clinical assessment information do not add any relevant information for the model for a significantly more accurate prediction: the mean RMSE value for FMA is 0.086 with MTA and 0.084 without MTA; the mean RMSE value for ARAT is 0.167 with MTA and 0.177 without MTA; the mean RMSE value for MI is 0.110 with MTA and 0.104 without MTA. Wilcoxon rank sum statistical test revealed a significant difference between the two conditions only for the ARAT clinical scale (for FMA p=0.519, z=-0.645, for ARAT p=0.032, z=2.145, and for MI p=0.066, z=-1.838). Looking at the figure 2b, it can be noted that ARAT scores have the widest distribution of cases in the analyzed patients population and ARAT predictions are the ones benefitting from the MTA in the input set, in terms of prediction accuracy. We believe, that the improvement in the ARAT clinical score outcome prediction when adding the MTA is occuring, because of the more spread distribution of the clinical score in all phases (T0, T1 and T2). In general, a more narrow distribution makes Random Forest Models to be limited in the prediction range, which make them rely heavily on limited patterns, while a more spread distribution provides more variance and therefore more patterns to learn. To the best of authors' knowledge there are no previous relevant studies that analyzed the influence of the mid-term assessment in a stroke rehabilitation outcome prediction task.

IV. Conclusions

The preliminary investigation on the mid-term assessment usage in the prediction of motor recovery of post-stroke patients revealed a little impact of the former in improving its accuracy. Prediction of the ARAT score only got a benefit from using the new information, but further analyses are needed. Models could be evaluated on fractions of the dataset which would be carefully generated so that the distribution in all time-points is more spread, and on other fractions of the dataset where the distribution is more narrow, highlighting the importance of the clinical time-points distribution.

V. Acknowledgements

This work has been funded by the MNESYS project (European Union- NextGenerationEU. PNRR – 'MNESYS – A Multiscale integrated approach to the study of the nervous system in health and disease'. Cod. PE_00000006) and the CO-AIDER project (put details here).

REFERENCES

[1] C. M. Stinear, W. D. Byblow, S. J. Ackerley, P. A. Barber, and M.-C. Smith, "Predicting recovery potential for individual stroke patients increases rehabilitation efficiency," *Stroke*, vol. 48, no. 4, pp. 1011–1019, 2017.

[2] M. Katan and A. Luft, "Global burden of stroke," in *Seminars in neurology*, vol. 38, no. 02. Thieme Medical Publishers, 2018, pp. 208–211.

[3] S. Campagnini, C. Arienti, M. Patrini, P. Liuzzi, A. Mannini, and M. C. Carrozza, "Machine learning methods for functional recovery prediction and prognosis in post-stroke rehabilitation: a systematic review," *Journal of NeuroEngineering and Rehabilitation*, vol. 19, no. 1, p. 54, 2022.

[4] C. M. Stinear and N. S. Ward, "How useful is imaging in predicting outcomes in stroke rehabilitation?" *International journal of stroke*, vol. 8, no. 1, pp. 33–37, 2013.

[5] C. M. Stinear, M.-C. Smith, and W. D. Byblow, "Prediction tools for stroke rehabilitation," *Stroke*, vol. 50, no. 11, pp. 3314–3322, 2019.

[6] C. Camardella, G. Cappiello, Z. Curto, M. Germanotta, I. Aprile, S. Mazzoleni, A. Scoglio, and A. Frisoli, "A random tree forest decision support system to personalize upper extremity robot-assisted rehabilitation in stroke: a pilot study," in *2022 International Conference on Rehabilitation Robotics (ICORR)*. IEEE, 2022, pp. 1–6.

[7] C. Camardella, M. Germanotta, I. Aprile, G. Cappiello, Z. Curto, A. Scoglio, S. Mazzoleni, and A. Frisoli, "A decision support system to provide an ongoing prediction of robot-assisted rehabilitation outcome in stroke survivors," in *2023 International Conference on Rehabilitation Robotics (ICORR)*. IEEE, 2023, pp. 1–6.

[8] I. Aprile, M. Germanotta, A. Cruciani, S. Loreti, C. Pecchioli, F. Cecchi, A. Montesano, S. Galeri, M. Diverio, C. Falsini, *et al.*, "Upper limb robotic rehabilitation after stroke: a multicenter, randomized clinical trial," *Journal of Neurologic Physical Therapy*, vol. 44, no. 1, pp. 3–14, 2020.

[9] Humanware. (2022) Motore, a robotic device for the rehabilitation of the upper limbs. [Online]. Available: https://www.hmw.it/en/robotic-neurorehabilitation/motore/

[10] L. Breiman, "Random forests," *Machine learning*, vol. 45, pp. 5–32, 2001.