Identifying Falls-Prone Older Adults Using Smartphone Sensing : A Machine Learning Approach

by

Cyrille MVOMO

BSc. Université Paris Cité

Research Internship / 3 months (Sept. 22 – Feb. 23) / MakerLab (INSERM UMR U1284)

Supervised by

Dr. Kevin LHOSTE

PhD.

Submitted

in Partial Fulfillment of the Requirements for the degree of

Master of Science in Sciences, Technologies, Health

(AIRE – Digital Sciences)

at

Université Paris Cité – Learning Planet Institute

January 2023

Table of Content

Identifying Falls-Prone Older Adults Using Smartphone Sensing : A Machine Learning Approach

by

Cyrille MVOMO

Submitted on January 25, 2023, in partial fulfillment

of the requirements for the degree of

Master of Science in Sciences, Technologies, Health

(AIRE – Digital Sciences)

Abstract *(Summary)*

One in three people over 65 falls at least once a year, resulting in 38 mil disability-adjusted life years lost each year. Considering our rapidly aging society, it has become crucial to develop accurate tools that allow therapists and clinicians to objectively identify and monitor falls-prone older adults. Despite the large body of technologies and literature in the field, we do not yet fully address the need for a system that would overcome both the lack of accessibility and transparency of current technologies and the lack of objectivity of conventional clinical tests, particularly when it comes to detect first-time fallers.

In this internship report, the objective was to conceptualize, design and develop a novel faller identification system. It consists in the first prototype of a mobile application that combines the use of classical threshold-based methods with machine learning, to help caregivers detect falls-prone older adults by supplementing conventional clinical assessment with functional test objectively measured using mobile sensing.

This report provides novel insight toward a more reliable and objective fall risk assessment paradigm using widely spread technology, especially when state-of-the-art devices are not available. Accordingly, next steps would include further developing and validating the mobile application for its deployment in clinical routine.

Keywords – *Fall Risk, Machine Learning, Smartphone Sensing, Older Adults, Threshold-Based, UX/UI, mHealth*

General Context

This internship took place at the MakerLab. The MakerLab is a place of innovation where frugal - reproducible and open - solutions to environmental and societal problems are imagined, prototyped and manufactured. It is a "FabLab" that allows students and collaborators of Learning Planet Institute to express their creativity by developing projects that meet the sustainable development goals.

The Learning Planet Institute is a non-profit organization linked to the Université Paris Cité and INSERM, whose mission is to explore, experiment and share new ways of learning and cooperating in order to address the needs of youth and the planet.

Although this project was a personal project that took place in my usual teaching place (i.e., no integration necessary), it was supervised by Dr. *Kevin LHOSTE*, robotics specialist and manager of the MakerLab, and addresses the SGD "good health and well-being".

INTRODUCTION

(Specific Context)

Falls are the second leading cause of unintentional injury deaths worldwide[1]. They are defined as "events that cause a person to rest inadvertently on the ground, floor, or any other lower level"[1]. While most healthy people take for granted the ability to maintain balance while standing or walking, the situation is quite different for our seniors over 60[1]. Indeed, this phenomenon is explained by the age-related degradation of the ability to cope with the mechanically unstable nature of antigravity control, and its need for integration of multiple sensory information coupled with the coordination of multiple muscles and joints unconsciously controlled by the brain[2].

The classical solution to help patients with balance disorders is to evaluate them using conventional clinical assessments in order to early detect them, and prescribe relevant therapeutic prevention programs, combined with a long-term monitoring. These assessments, which aim to consider the multifactorial nature of falls, are generally based on the retrospective fall history (e.g., number of falls in the past year), clinical questionnaires, gait analysis and clinical posturography. However, they suffer from a low level of objectivity expressed by an operator-dependent reliability, which prevents them from detecting first-time fallers.

With the democratization of new information and communication technologies in healthcare, and the rise of evidence-based medicine, many innovative screening systems have been developed to complement the conventional assessment by allowing the objective identification of fall-prone older adults. However, most of these systems remain poorly deployed in clinical routine for reasons mainly related to their lack of reliability and accessibility, but also their poor implementation, practicality, or usability. Caregivers' rejection of the adoption of these tools can be explained by their poor transparency, relevance (i.e., consideration of the multifactorial nature of falls) or accuracy (i.e., reliability), their often-expensive price or installation requirements (i.e., accessibility), their difficulty in adapting to the functioning of health centers and clinical practice (i.e., implementation and practicality), or the need for external technical expertise (i.e., usability). Therefore, an efficient faller identification technology that allows to objectify conventional clinical assessments without all the aforementioned limitations of current state-of-the-art systems is yet to be developed.

In this internship report, we expand on previous work focusing in using data-driven approaches to build a comprehensive faller risk identification framework and combine it with conventional clinical assessments into a novel mobile health application called *MakerRehab©*[3-9].

Thus, in the study of *Doheny et al. (2012)*, 110 healthy older adults above 60 wearing an accelerometer placed on the hip, were instructed “to stand as still as possible for 35 seconds” in the comfortable and semi-tandem stances[8]. In this study, the participants were classified as fallers and non-fallers based on their retrospective fall history. With this article, the authors have shown the ability of triaxial accelerometry placed on the hips to discriminate fallers from non-fallers by measuring postural sway. *Amick et al. (2015)* and *Nishiguchi et al. (2012)*, extended these findings by showing the ability of smartphone's built-in triaxial accelerometer-based mobile applications to measure gait and balance, in concurrent validity studies conducted in healthy young adults[3,4]. Finally, these works have been completed by *Cerrito et al. (2014)* in another concurrent validity study conducted in healthy seniors demonstrating the reliability of the hip placement for its ability to approximate the position of the body center of mass[9]. Despite these avenues, these works did not focus on the identification of fallers but rather on the objective quantification of balance during functional test using mobile health (mHealth) technologies. Data-driven approaches have filled this gap by allowing the identification of fallers from sensor-assisted functional tests, but also from conventional clinical assessments. Using threshold-based methods and/or machine learning techniques, *Fahimi et al. (2021)* and *Shumway-Cook et al. (1997)* have successfully built models of faller screening[5,6]. These approaches have shown their robustness due to their ability to identify both prospective (i.e., first-time fallers) and retrospective fallers, but also to take into consideration the multifactorial nature of the falls that is crucial to the management in clinical routine[5]. However, their deployment in clinical practice is still in its infancy.

In view of the current literature, it appears obvious that there is a need for a system implementing these data-driven approaches to identify fallers in clinical routine, by combining them with approaches quantifying balance using mHealth technologies. Therefore, this study aimed to (i) conceptualize, (ii) design and (iii) develop the first prototype of a mobile application for caregivers to identify fallers and quantify balance performance for diagnostic or follow-up purposes, using a combination of threshold-based methods and machine learning, relying on conventional clinical assessments and smartphone sensing-based functional test.

Accordingly, we hypothesized *MakerRehab©* (i) to be robust in identifying faller and quantifying fall risk, (ii) to follow the user-centered design framework for mHealth technologies, and (iii) to align with the recommendations for the development of mHealth applications[10,11].

Are we able to produce a mHealth technology that addresses the challenges of current practices and devices for the screening and monitoring of fallers?

MATERIAL & METHODS

(Data - Methods - Implementation)

This section describes the material and methods used to conceptualize, design and develop *MakerRehab©*. All the codes and dataset are available online : https://github.com/cyrillemvomo/FallRiskAPP .

I. *Material*

Data —— In this study, we extracted our data from a publicly available database[12]. The database can be accessed online (https://doi.org/10.7717/peerj.2648). The content of the database and the way in which the data were collected were detailed by the authors in *Santos et al. (2016)*[12]. Data were collected from 163 men and women of ages ranging from 18 to 86 years with varying health conditions. The data collection process consisted in a 1-2 hour single session per participant. A session consisted in the performance of a conventional clinical assessments of the risk of falling including the retrospective fall history (e.g., number of falls in the past year), clinical tests, questionnaires and clinical posturography.

For the clinical posturography, they were instructed to stand barefoot in a constrained position on a force platform as still as possible for 60 seconds with arms at the side of the body and repeat three times for four different test conditions while the force platform measured the participant's displacement of the plantar center of pressure (COP), force, and force moment in the x, y, z directions (except for the displacement of the COP) at a sampling frequency of 100 Hz. The four conditions were eyes-open on firm surface, and eyes-open on foam surface, eyes-closed on firm surface, eyes-closed on foam surface [*figure 1*].

As each participants did the 60 seconds-long standing balance test 12 times (i.e., 3 times for each 4

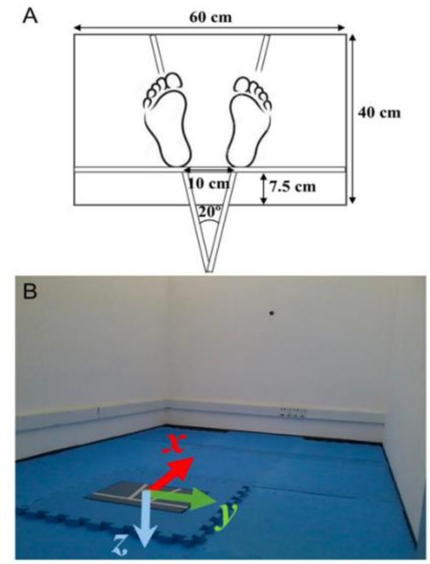


Figure 1: *Clinical Posturography measurement set-up from Santos et al. (2016)*

conditions), the database consisted in 1632 text files. Each text file contained 6000 rows and 9 columns representing the time (i.e., 1 column), the tri-axial force smoothed with a 10 Hz 4th order zero lag low-pass Butterworth filter and force moment (i.e., 6 columns) and the displacement of the COP on the x and y axis (i.e., 2 columns).

In addition to these text files, the database contained an additional csv file containing general information (e.g., fall history, sex, age, BMI, ID, medication, …) about the participants and their performance on the other clinical tests. This additional file a total of 53 columns and 1632 rows. These clinical tests included the Falls Efficacy Scale (FES) for the assessment of fear of falling, International Physical Activity Questionnaire (IPAQ) for the evaluation of physical condition, and Mini Balance Evaluation Systems Test (mini-BEST) for the subjective evaluation of balance performance. All these tests are valid and widespread in clinical routine, as part of a fall risk assessment.

With its 11 580 000 of rows of data, this database is distinguished by the level and quality of information contained. Having opened the possibility to researchers to answer many questions and apply data-driven approaches which often requires a large amount of data[7]. The construction of this database has been ethically approved by a committee of the *Federal University of ABC, São Bernardo do Campo, SP, Brazil*. The participants consented to the sharing of their data and were anonymized by the authors. However, the database can only be used for research purposes.

Technical Environment —— Concerning the conceptualization of *MakerRehab©*, both the extraction, clea-

-ning exploration of the training data, as well as the construction and evaluation of the predictive model have been conducted on a MacBook Air M1 2020 computer (*Apple silicon, Apple Inc. USA, MacOS version 13.1*) and implemented in Python 3.10.4 using Scikit-learn library version 1.1.1, Pandas library version 1.3.5, Matplotlib library version 3.5.1, Plotly library version 5.8.0, Seaborn library version 0.11.2, Scipy library version 1.8.0, and NumPy library version 1.21.5.

For the design part, the User-Centred Design Framework for mHealth technologies (UCDFmH) by *Farao et al. (2020)* was applied[10].

Finally, for the development, Flutter version 3.3.10, the open-source UI (i.e., User Interface) cross-platform software development framework, from Google (*Google LLC. USA*) was used with Dart version 2.18.6 along with Python and its Flask library version 2.2.2 in the backend, and the Guidelines for the Development of mHealth applications (GDmH) from *Chatzipavlou et al. (2016)* were followed[11].

Hardware —— The hardware part of *MakerRehab©*, consisting in the smartphone attachment system for the functional tests, is composed of two parts: A low-cost velcro belt available online, and a 3D printed holder (*Ultimaker 3 3D printer, Ultimaker BV, ND*) modelled using Fusion360 version 2.0 (*Autodesk Inc. USA*) [*figure 2*].

II. *Methods*

Conceptualization —— Once the literature was reviewed and the objectives were clearly defined, the conceptualization phase of *MakerRehab©* consisted in building, cleaning and exploring our dataset before implementing our classifiers.

Therefore, the preparation of our dataset has been summarized in figure 3. The first step was to collect the data by filtering the database in order to retain only those subjects who met the inclusion criteria. Thus, to be retained, the subjects had to be free of any pathology that could directly affect balance, to have no missing data and to have completed the 12 required measurements.

Once collected, the subjects were divided into "fallers" or "non-fallers" according to the information available in the database. The "fallers" corresponded to subjects over 60 years old, having fallen at least once in the last 12 months and who had: either a score lower than or equal to 16/28 in the miniBEST test (i.e., balance disorder), or a "high concern" about the fear of falling, or a completion time of tasks A and B of the Trail Making test (i.e., executive functions test) of 5 minutes

Figure 2: *3D printed holder (left) and its velcro belt (right)*

(i.e., executive functions disorder), or a low physical activity[12]. The "non-fallers" corresponded to elderly or young subjects who had not fallen in the last 12 months and who had either a score higher than 16/28 in the miniBEST test, or a "moderate concern" or a "low concern" about the fear of falling, or a completion time of tasks A and B of the Trail Making test of less than 5 minutes, or a "high" or "moderate" physical activity.

At the end of this step, the subjects were given an ID. For each of the 12 trials of each subject, the mediolateral (i.e., y-axis) and anteroposterior (i.e., x-axis) displacement of the body center of mass (COM) was estimated from the COP displacement signals with the function described by *Caron et al. (1997)* and implemented in Python by Professor *Marcos Duarte* (*Federal University of ABC, São Bernardo do Campo, SP, Brazil)*[12] :
https://nbviewer.org/github/BMClab/BMC/blob/master/notebooks/IP_Model.ipynb.

From these newly calculated COM signals, 9 functional parameters, which have been identified in the literature as being able to discriminate fallers from non-fallers, were calculated for each of the 4 conditions during the 3 trials per condition performed by each subject (i.e., a total of 36 functional parameters) [8]. These functional parameters were summarized in table 1 and constituted the first part of the dataset.

Finally, 10 critical information well known as playing a role in the risk of falling and balance performances were extracted from the csv file contained in the database, then attached to the first part of the dataset containing a total of 47 columns (i.e., 36 for the functional parameters, 10 for the critical information, and 1 for the class) and 393 rows (i.e., 3 trials for each 131 subjects). The 10 critical information were the following ones: Age, Age Group, Gender, Height, Weight, BMI, Number of falls during the past 12 months, miniBEST test performance, FES scale score, and IPAQ questionnaire score.

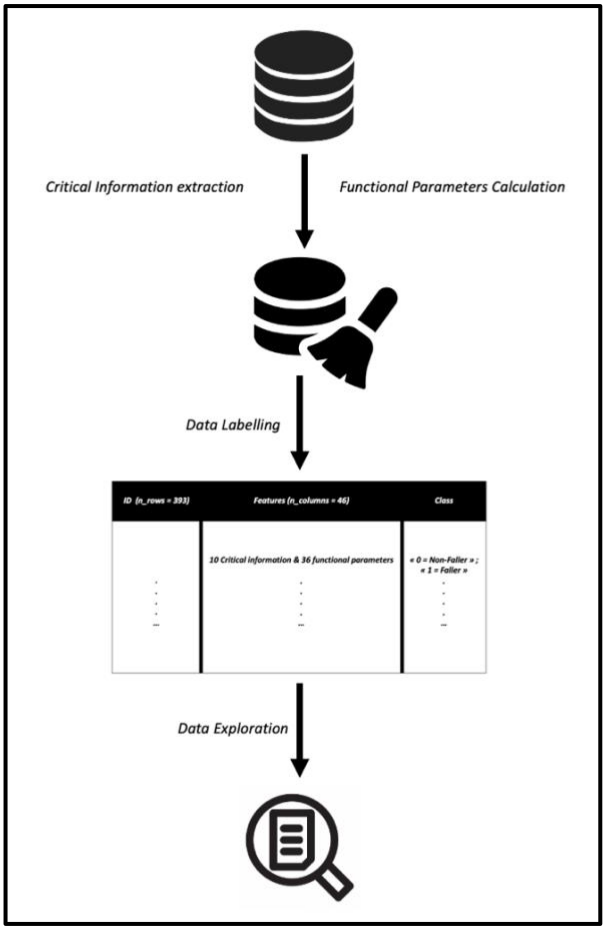


Figure 3: *Data preparation pipeline*

Once our dataset was built, its cleaning consisted in labelling the categorical features and the class using a manual encoding technique. Thus, regarding the features, the age groups "old" and "young" respectively received a 1 and a 0, the gender "female" and "male" respectively received a 0 and a 1, and the scores "high", "moderate", and "low" to the IPAQ questionnaire respectively received a 3, a 2, and a 1. As for the class, the "fallers" received a 1 and the "non-fallers" received a 0.

As a conclusion to this conceptualization phase, an exploratory data analysis was performed in order to determine the exact composition of our dataset and identify the non-causal relationships between the features and the class.

In this exploratory data analysis summarized in Figure 4, Spearman correlation analysis revealed strong positive linear correlations between functional parameters for each condition ($r^2 > 0.7$) but weak between conditions and with class ($0 < r^2 < 0.5$). For the critical information, only the number of falls was positively linearly correlated with the class ($r^2 > 0.7$), the other information thus presenting positive ($r^2 > 0$)

Table 1: *Functional Parameters*

Functional Parameter	Formula	Definition
Area (m^2)	$Area = \pi \cdot \lvert e' \rvert \cdot \lvert e' \rvert$ by Wollseifen et al. 2011[13]	Area of the displacement of the body center of mass
Range AP (m)	$Range = max[x] - min[x]$ with x representing the COM medio-lateral/antero-posterior displacement array	Amplitude of displacement of the body center of mass in the antero-posterior direction (i.e., x axis)
Range ML (m)		Amplitude of displacement of the body center of mass in the medio-lateral direction (i.e., y axis)
Sway Length AP (m)	$Total\ path = \sum_{i=1}^{N-1} \lvert x_{i+i} - x_i \rvert$ with x representing the COM medio-lateral/antero-posterior displacement array and N representing the length of x	Length of displacement of the body center of mass in the antero-posterior direction (i.e., x axis)
Sway Length ML (m)		Length of displacement of the body center of mass in the medio-lateral direction (i.e., y axis)
Mean Velocity AP (m/s)	$Mean\ speed = \frac{1}{T} \sum_{i=1}^{N-1} \lvert x_{i+i} - x_i \rvert$ with x representing the COM medio-lateral/antero-posterior displacement array and N representing the length of x and T representing the time array	Mean velocity of the body center of mass in the antero-posterior direction (i.e., x axis)
Mean Velocity ML (m/s)		Mean velocity of the body center of mass in the medio-lateral direction (i.e., y axis)
RMS Acceleration ML (g)	$RMS = \sqrt{\frac{1}{N} \sum_{i=1}^{N} x_i^2}$ with x representing the COM acceleration array and N representing the length of x	Root mean square acceleration of the body center of mass in the antero-posterior direction (i.e., x axis)
RMS Acceleration AP (g)		Root mean square acceleration of the body center of mass in the medio-lateral direction (i.e., y axis)

Table 2: *Critical Information (mean +- std) for fallers and non-fallers, and p-value of the differences between groups*

	Fallers (mean+/- std - n=36)	Non-Fallers (mean+/- std - n=357)	P-value
Age	71.04167 (+/- 7.27228)	47.80112 (+/- 22.96759)	0.00000
Height	156.94167 (+/- 4.33598)	163.1521 (+/- 10.05898)	0.00029
Weight	62.1375 (+/- 7.6607)	63.29958 (+/- 7.91038)	0.40123
BMI	25.16807 (+/- 2.42019)	23.89607 (+/- 3.24504)	0.02298
Falls12m	1.08333 (+/- 0.27639)	0.0 (+/- 0.0)	0.00000
Best_T	17.33333 (+/- 4.30762)	21.69748 (+/- 3.78976)	0.00000
FES_T	9.75 (+/- 2.77263)	10.4958 (+/- 3.20942)	0.18065
AgeGroup(0=young|1=old)	1.0 (+/- 0.0)	0.46218 (+/- 0.49857)	0.00000
Gender(0=F|1=M)	0.08333 (+/- 0.27639)	0.31933 (+/- 0.46622)	0.00309
IPAQ_S	1.91667 (+/- 0.7592)	2.19328 (+/- 0.72503)	0.03087

or negative ($r^2 < 0$) linear correlations generally weak ($0 < r^2 < 0.5$ or $0 > r^2 > -0.5$) to moderate ($0.5 < r^2 < 0.7$ or $-0.5 > r^2 > -0.7$) between them and weak with the class, except for age with the age group ($r^2 > 0.7$), and the performance at the miniBEST test with the age and the age group ($r^2 < -0.7$).

The critical information data have been summarized in Table 2 and a One-way analysis of variance (ANOVA) was used to examine the difference in each measure between fallers and non-fallers.

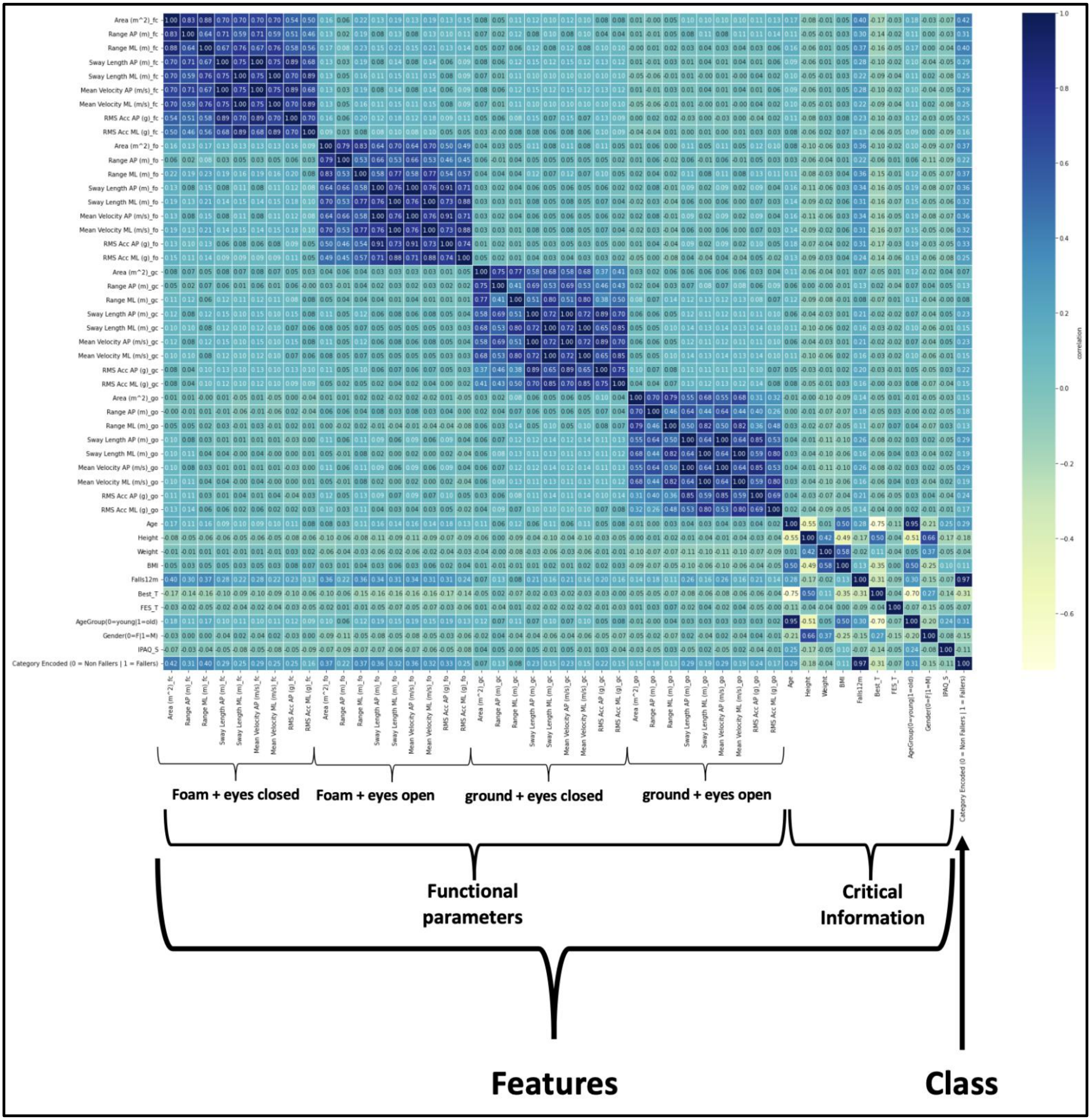


Figure 4: *Exploratory Data analysis (Spearman correlation heatmap)*

Implementation —— As soon as the conceptualization phase was completed, the implementation phase aimed at developing our classifiers and integrating them into the Flutter application following the *UCDFmH* and *GDmH*.

Accordingly, for the supervised learning task, we sought to implement a binary classifier capable of detecting fallers and quantifying balance performance, considering the multifactorial nature of falls and performing at least as well as similar models in the literature[5]. Therefore, we wanted an explainable model, using all the information of the dataset. Hence, we could neither use deep learning methods employing inexplicable neural networks, nor remove features, drop outliers or use data augmentation techniques.

Consequently, *Fahimi et al. (2021)* approach was heavily relied upon and adapted to our modeling task consisting in mapping the 46 features (i.e., input) to their respective class (i.e., output)[5]. Using various machine learning algorithms, we performed a repeated 10-fold cross-validation (CV) with 5 repetitions. In each fold, we trained the model on 75% of the data and tested on the remaining 25%.

Three machine learning algorithms - known for their performance in this kind of task[5] - were tried and compared. These were logistic regression (LR), k-nearest neighbors (KNN), and support vector machine (SVM). For each algorithm, the following pipeline was followed: (i) pre-training of the model without hyperparameter optimization; (ii) pre-evaluation of the performance on the test set with the area under the ROC (AUROC), accuracy, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV)[5]; (iii) pre-assessment of the learning quality by comparing the performance on the training and test sets with the learning curves; (iv) regularization by optimizing the hyperparameters using GridSearch with cross-validation; (v) re-training the model with the optimal hyperparameters; (vi) re-evaluation of the performance on the test set with AUROC, accuracy, sensitivity, specificity, PPV, and NPV; (vii) re-evaluation of the learning quality by comparing the performance on the training and test sets with the learning curves; (viii) identification of the model with the best AUROC[14]; (ix) statistical analysis of the differences in AUROC of each of the models with Kruskal-Wallis and post-hoc comparison and a level of statistical significance fixed at 0.05; (x) evaluation of how the model works with a feature importance analysis.

With an AUROC score of 98.89%, LR was the best model for mapping each participant's functional parameters and critical information to their fall risk. In addition to its AUROC score, LR presented the best accuracy, sensitivity and PPV [table 3]. Despite the fact that Kruskal-Wallis revealed that the median of the AUROC scores of each model were significantly different (p-value <0.05), post hoc comparisons did not reveal any statistical difference between LR and SVM AUROC scores (p-value > 0.05) meaning that LR wasn't performing significantly better than SVM [table 3, 4]. Since LR's performance was similar to or better than that of other models in the literature for the same type of task or with the same dataset, this model was selected for *MakerRehab©*. The whole pipeline used for LR has been summarized in figure 5, the feature importance analysis in figure 6A, the AUROC curve in figure 6B, LR and SVM tuned-hyperparameters in figure 7A, and LR final learning curves in figure 7B. As shown, the final model was learning correctly [figure 7B]. However, surprisingly, it gave almost no importance to functional parameters compared to critical information such as the number of falls in the last 12 months [figure 6A]. However, this fact did not prevent us from retaining LR as a predictive model.

In addition to the machine learning-based classification, a threshold-based classification has been implemented. Subsequently, *Doheny et al. (2012)*

Figure 5: *Machine Learning Modelling Flowchart*

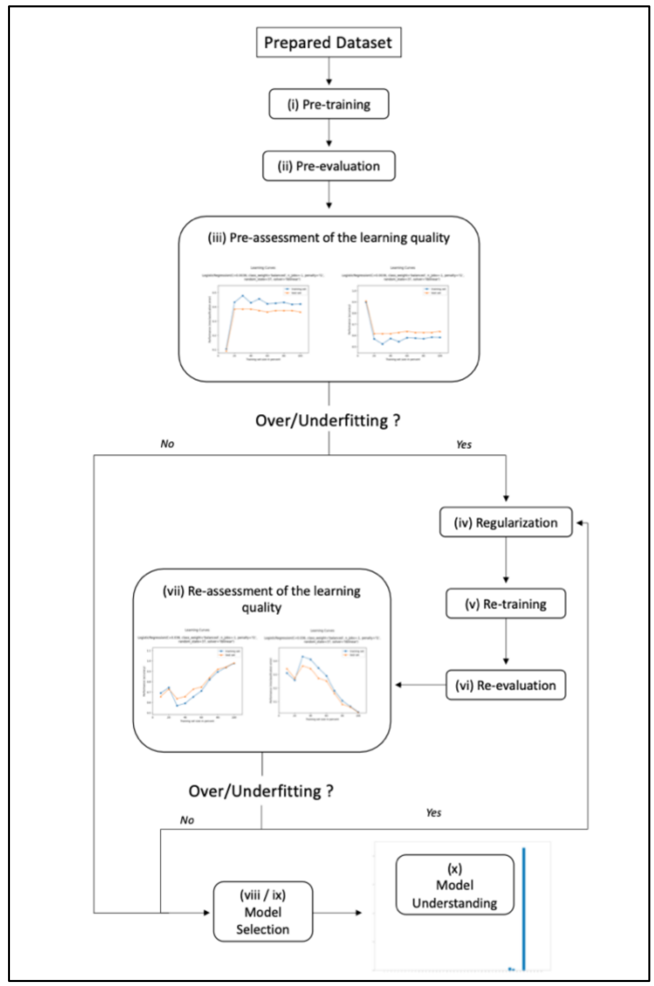


Table 3: *Models Performances (in %)*

Algorithm	AUROC	Accuracy	Specificity	Sensitivity/Recall	PPV/Precision	NPV
LR	98.89	97.98	97.78	100.00	81.82	97.78
KNN	71.67	93.94	98.89	44.44	80.00	98.89
SVM	96.67	93.94	93.33	100.00	60.00	93.33

Table 4: *Statistical Analysis of AUROC (post hoc – p-value)*

	LR	SVM	KNN
LR	1.000000e+00	1.000000e+00	1.372857e-19
SVM	1.000000e+00	1.000000e+00	4.436229e-22
KNN	1.372857e-19	4.436229e-22	1.000000e+00

Figure 6: *Feature Importance Analysis (A) rounded with 2 decimals and LR AUROC Curve (B)*

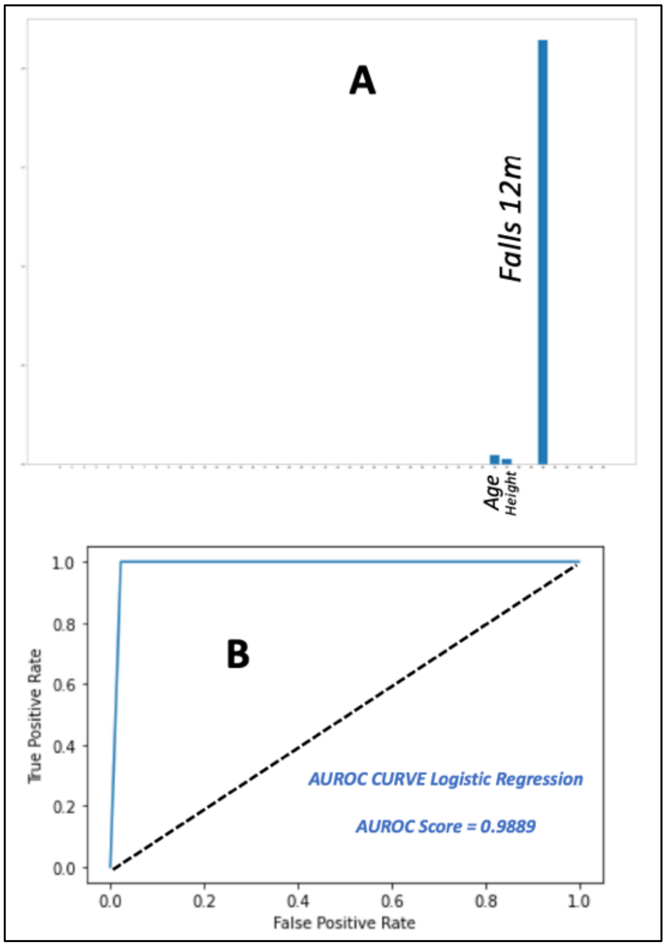
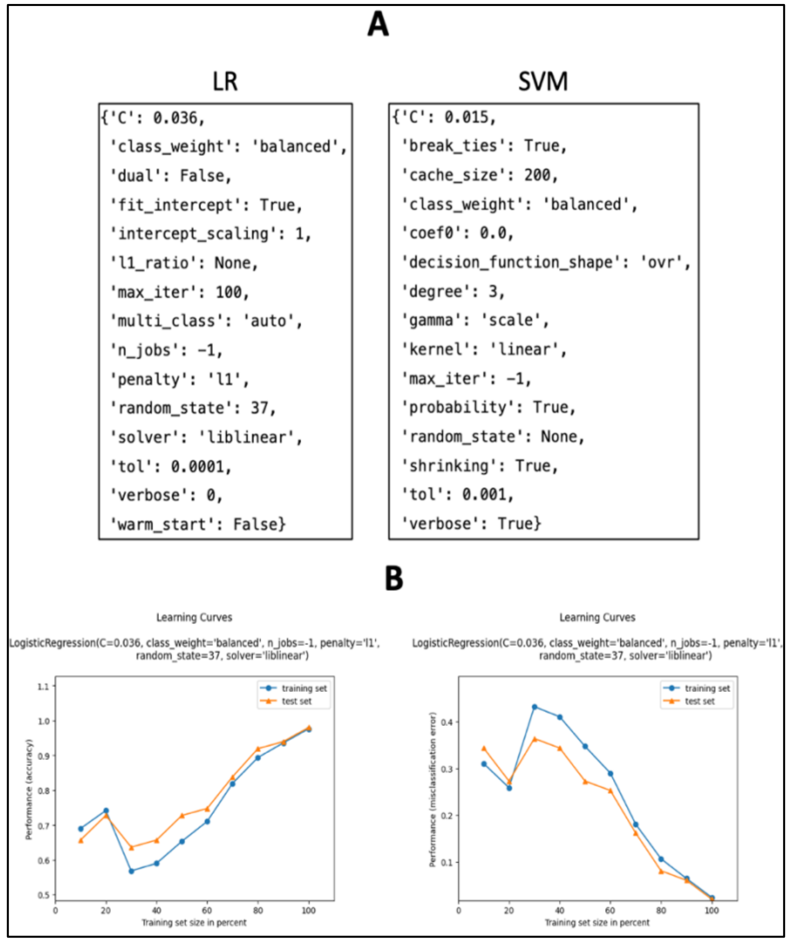


Figure 7: *LR and SVM tuned-hyperparameters (A), and LR final learning curves (B)*



results were associated with the data from our dataset in order to find the optimal thresholds discriminating fallers from non-fallers for each of the 36 functional parameters used in our case[8]. While the thresholds of the features recognized as being the most discriminating were based on the literature, those being less discriminating were based on our dataset[8]. These thresholds can be observed from Table 5.

Once our classifiers were ready to be deployed, it was time to think about the design of *MakerRehab©*. For this purpose, we used the *UCDFmH* of *Farao et al. (2020)* [10]. Based on the combination of the principles of "Design Thinking" and the "Information Systems Research framework", the *UCDFmH* consists of a succession of steps allowing both to optimize the engagement of end-users, and a rapid and low-cost development responding to contextual challenges and needs[10]. These attributes are central to the development of mHealth technologies such as *MakerRehab©*, therefore we opted to implement the *UCDFmH*. Thus, the framework consists of 3 cycles: the relevance cycle, the design cycle, and the rigour cycle [figure 8].

The relevance cycle consisted in identifying the

Table 5: *Functional Parameters Pathological Thresholds from our dataset and Doheny et al. 2012 (in grey)* [8]

Functional Parameters	Pathological Threshold (= mean(faller) + std(faller)) → Faller ≥ Pathological Threshold > Non-Faller
Area (m^2)_fc	0.00252
Range AP (m)_fc	0.06047
Range ML (m)_fc	0.0566
Sway Length AP (m)_fc	0.44116
Sway Length ML (m)_fc	0.36988
Mean Velocity AP (m/s)_fc	0.0147
Mean Velocity ML (m/s)_fc	0.01233
RMS Acc AP (g)_fc	0.02312
RMS Acc ML (g)_fc	0.01904
Area (m^2)_fo	0.00192
Range AP (m)_fo	0.04956
Range ML (m)_fo	0.04922
Sway Length AP (m)_fo	0.38014
Sway Length ML (m)_fo	0.31924
Mean Velocity AP (m/s)_fo	0.01267
Mean Velocity ML (m/s)_fo	0.01064
RMS Acc AP (g)_fo	0.02127
RMS Acc ML (g)_fo	0.01791
Area (m^2)_gc	0.0002
Range AP (m)_gc	0.02167
Range ML (m)_gc	0.01314
Sway Length AP (m)_gc	0.13618
Sway Length ML (m)_gc	0.08676
Mean Velocity AP (m/s)_gc	0.00454
Mean Velocity ML (m/s)_gc	0.0029
RMS Acc AP (g)_gc	0.01376
RMS Acc ML (g)_gc	0.01089
Range AP (m)_go	0.019
Range ML (m)_go	0.0127
Sway Length AP (m)_go	0.19
Sway Length ML (m)_go	0.126
Mean Velocity AP (m/s)_go	0.00991
Mean Velocity ML (m/s)_go	0.00630
RMS Acc AP (g)_go	0.05
RMS Acc ML (g)_go	0.03

end-users' environment with 2 modes: the emphasize mode, and the define mode. For this, the recommendation of the French High Health Authority (HAS) was followed in order to understand exactly the context of the end-users and the problems they may face during the clinical examination of the risk of falling[15]. From this recommendation, 2 following characteristics that *MakerRehab©* should incorporate were retained:

- The assessment is multifactorial
- The assessment is based on valid and standardized tests

Following the relevance cycle, the design cycle involved using the ideate and prototype modes. To do this, ideas from technologies similar to *MakerRehab©* were generated and modeled on a personal notebook[4] [figure 9A]. Thanks to these sketches, a low-quality pre-prototype of *MakerRehab©* could be made and continuously improved [figure 9B].

Figure 8: *UCDFmH from Farao et al. 2020*[10]

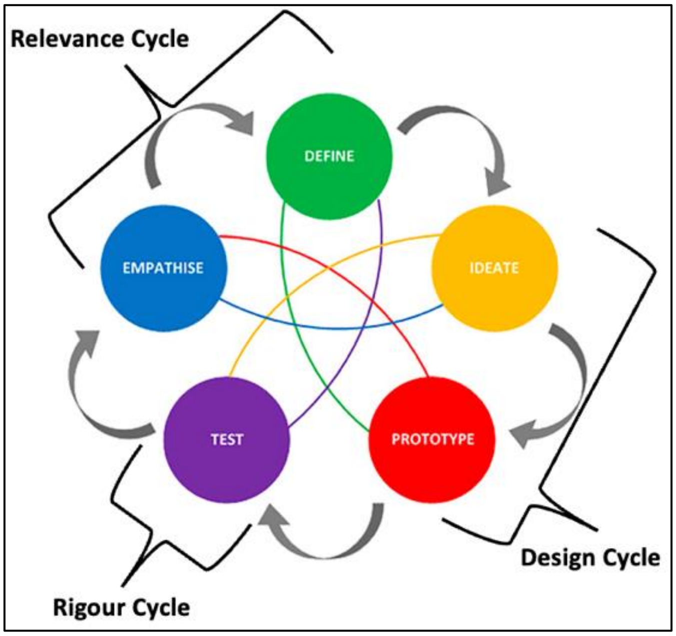
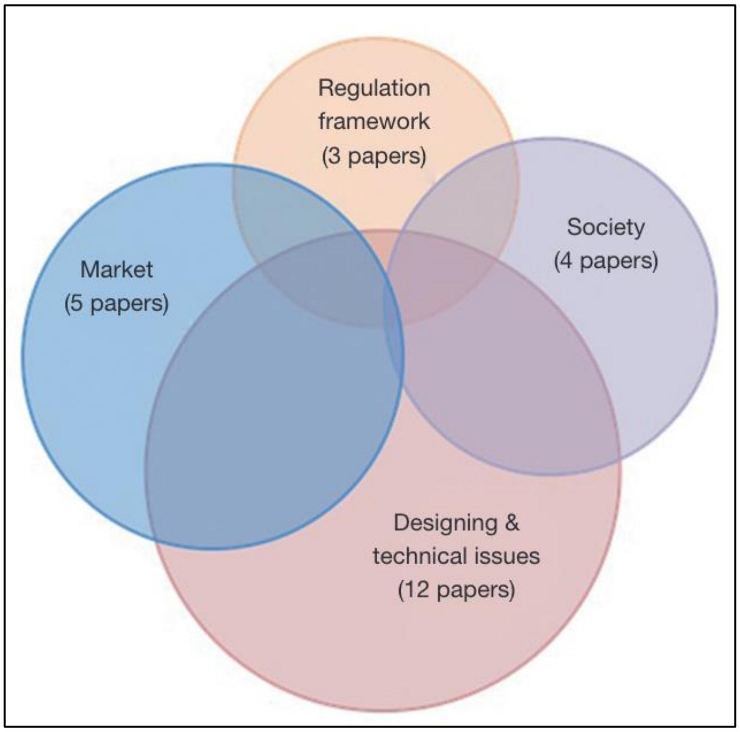


Figure 9: *Sway Medical Balance App© (A)*[4] *and Pre-prototype of MakerRehab© (B)*

Finally, the rigour cycle involved the test mode where *MakerRehab©* was first reviewed and then evaluated against the two features defined in the relevance cycle, and the 4 pillars of the *GDmH*[11] [figure 10]. These pillars included consideration of issues related to:

- Regulation framework concerning the consideration of regulatory issues that *MakerRehab©* will have to face
- Society regarding the understanding of ethical and societal issues inherent to the development of *MakerRehab©* (e.g., capacity to deploy in low- and middle-income countries)
- Market concerning the understanding of the market necessary for the deployment of *MakerRehab©*
- Designing and technical issues that need to be addressed.

Figure 10 : *GDmH from Chatzipavlou et al. (2016)*[11]



RESULTS

This section describes and discuss the final version *MakerRehab©* and the results of the rigour cycle. All the codes are available online: https://github.com/cyrillemvomo/FallRiskAPP.

The final version of *MakerRehab©* was composed of 2 parts including (i) a *Main Page, (ii)* a *Home page*.

Main Page —— As shown in figure 11A, the main page simply constituted the introduction page of the application and allows the user to access the home page by pressing "*Get Started*".

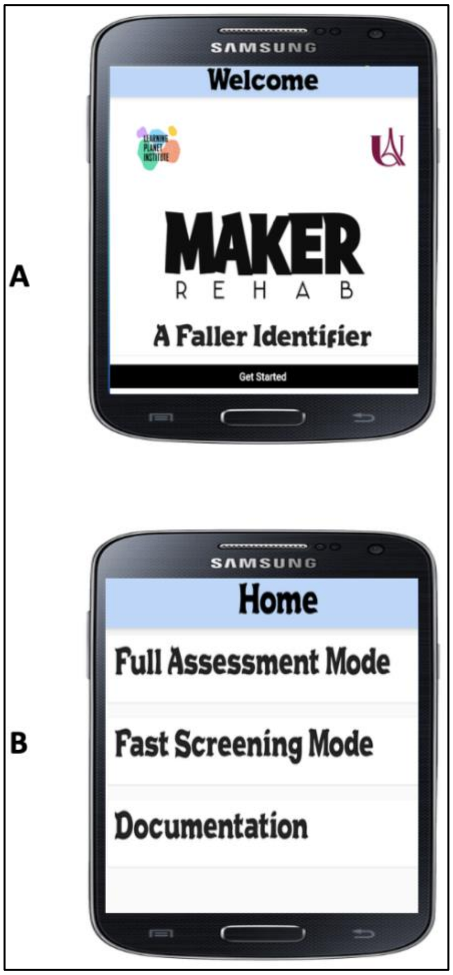
Home Page —— The home page is divided into three parts that make up the menu of *MakerRehab©*, with:

- The *Full Assessment Mode*
- *Fast Screening Mode*
- and the *Documentation*

While the *Full Assessment* and *Fast Screening* modes are the two modes of fall risk assessment in *MakerRehab©*, by clicking on "Documentation", the user will have access to this report providing transparency on how the application works and has been built [figure 11B].

The *Full Assessment* and *Fast Screening* modes differ in that, while the *Full Assessment* mode is a comprehensive and multifactorial fall risk assessment

Figure 11: *Main (A) and Home (B) pages of MakerRehab©*



including conventional clinical tests and smartphone sensing-assisted functional tests, the *Fast Screening* mode is a rapid screening including only smartphone sensing-assisted functional tests.

Full Assessment Mode **——** The different steps of the *Full Assessment Mode* have been summarized in annexe 3. It begins with an instruction page that gives the caregiver an overview of the entire workflow by illustrating the 4 steps of the mode, as well as an informed consent to obtain patient consent, and to let the patient know his or her rights and how the data he or she is providing will be handled. Its 4 steps include: (i) general information where the practitioner can fill in patient-specific information such as weight, height, annual fall history, age, gender at birth, among others; (ii) clinical tests and questionnaires consisting of conventional clinical tests allowing the practitioner to perform a subjective assessment of the patient's physical condition using the IPAQ questionnaire, balance performance with the mini-BEST test, and fear of falling with the FES-I scale; (iii) the *MakerRehab©* Balance Test consisting of a static barefoot standing balance test where balance is measured under 4 different conditions using the smartphone accelerometer (placed on the lower back to collect the COM movement data): eyes closed on a firm floor (visual sensory disturbance), eyes closed on a compiling floor (visual and proprioceptive sensory disturbance), eyes open on a firm floor (no sensory disturbance), eyes open on a compiling floor (proprioceptive disturbance); (iv) and finally the clinical report containing all the information collected, the results of the different tests performed, the balance score and the label [figure 12].

Throughout the mode, each information entered by the user is saved. In addition to being saved, the collected data are transmitted to a local Python server that processes the signals collected by the accelerometer during the *MakerRehab©* Balance Test in order to obtain the displacement of the center of mass and thus calculate the different functional parameters, predict the risk of falling using the predictive model, and calculate the balance score.

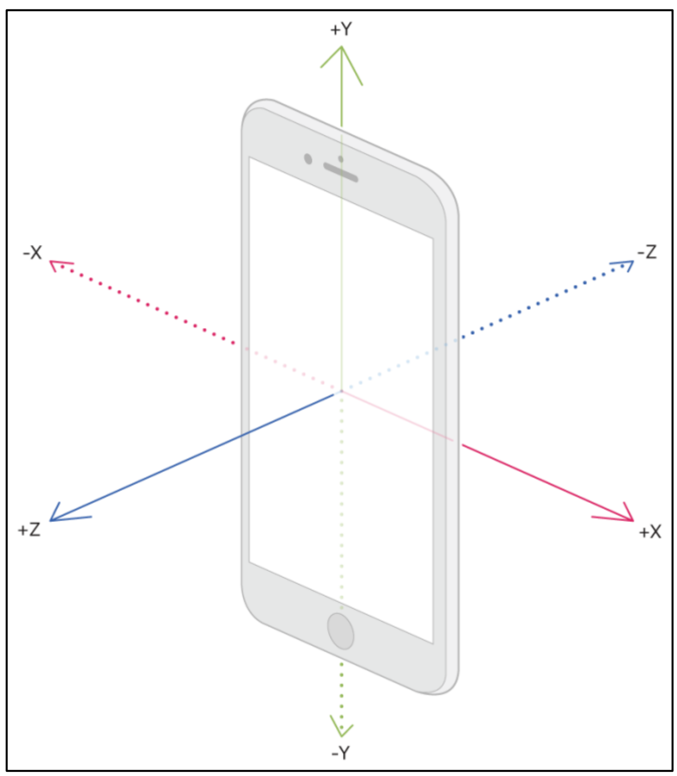
The signal processing was performed following the approach of Doheny et al. (2012) to avoid accelerometer drift using the trapezoidal method consisting in applying a 6th order Butterworth high-pass filter with zero phase shift on the acceleration signal at the optimal cutoff frequency calculated by residual analysis, first integrating the acceleration signal of the center of mass in order to get the velocity signal of the center of mass, apply again a 6th order Butterworth high-pass filter with zero phase shift on the velocity signal at the optimal cutoff frequency computed by residual analysis, and finally integrate the velocity signal of the center of mass in order to have the displacement signal necessary to compute the functional parameters [figure 13]. Once the signal has been processed and the 36 functional parameters have been calculated, the server joins them to the other data collected, thus constituting the 46 features necessary to predict the risk of falling and calculate the balance score.

Accordingly, the server then sends all these features as input to our predictive model and the predictive model returns both a "high" or "low" class corresponding to a high or low risk of being a faller, and a score out of 100 corresponding to the balance score based itself on the estimation of the probability to belong to the "faller" class. The score formula is therefore as follows:

$$\textit{Balance Score} = (\,1 - P_{estimate}(\mathbf{X})\,) \times 100$$

where $P_{estimate}(\mathbf{X})$ is the probability to belong to the "faller" class. In other words, the higher the balance score, the higher the balance performance.

Figure 12: *Smartphone's Tri-Axis Accelerometer used by MakerRehab©*



Fast Screening Mode **——** It begins with the same instruction page as the *Full Assessment Mode*. It organized in 3 steps including: (i) general information where the practitioner can fill in the same information as in the *Full Assessment Mode*; (ii) the *MakerRehab©* Balance Test; (iv) and the clinical report containing all the information collected, the results of the different tests performed, the balance score and the label.

Throughout this mode, each information entered by the user is also saved and transmitted to the same local Python server that processes the signals collected by the accelerometer during the *MakerRehab©* Balance Test in order to obtain the displacement of the center of mass and thus calculate the different functional parameters, predict the risk of falling, and calculate the balance score using the threshold-based classifier.

Accordingly, the server then first calculates the balance score using the following formula:

$$\text{Balance Score} = 1 - \left(\sum \text{Patient\_Functional\_Parameters} \times 100\right) / \sum \text{Pathological\_Thresholds}$$

where $\sum \text{Patient\_Functional\_Parameters}$ is the sum of all the 36 calculated functional parameters, and $\sum \text{Pathological\_Thresholds}$ is the sum of all the 36 defined pathological thresholds (i.e., the ones that discriminates fallers from non-fallers) [Table 5]. After calculating the score, if $\text{Balance Score} > 50$ if the classifier return “high” meaning a high fall risk, and if $\text{Balance Score} \leq 50$ the classifier return “low” meaning a low fall risk. Subsequently, the higher the balance score, the higher the balance performance.

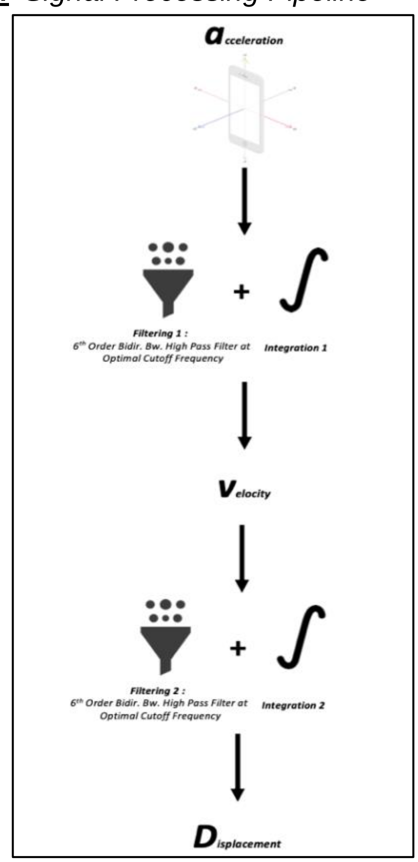
In both modes, when the functional parameters have been calculated, the label obtained and the score established, the server sends them, along with all the other features, to the application which displays them in the clinical report, adding graphs comparing the patient's performance in the *MakerRehab©* balance test to the pathological thresholds [Table 5]. At that precise moment, the user can either visualize it from the application (i.e., the clinical report) or extract it in pdf format. It should also be noted that the caregiver has access to a "help" button throughout the performance of the conventional clinical tests, giving him a detailed explanation of how the different tests should be performed, and he also has the possibility to extract the raw acceleration data along with the report.

Evaluation **——** The evaluation of *MakerRehab©* aimed at ensuring that the application met the 2 characteristics established during the relevance cycle and followed the 4 pillars of *GDmH*. Thus, by integrating the 3 conventional clinical tests in the full assessment mode, we met the need for an evaluation based on valid and standardized clinical tests. Moreover, by assessing both the psychological (e.g., fear of falling) and physical (e.g., *MakerRehab©* Balance test) aspects, we also met the need for a multifactorial assessment.

As for the <u>Regulatory Framework</u>: *MakerRehab©* will have to demonstrate its value to the competent entities for its future deployment (e.g., FDA or European Commission). For the moment, being still at the prototype stage, a thorough evaluation of the regulatory aspect cannot be done.

Regarding the <u>Market</u>: The target user group and diagnostic nature has been clearly identified. It is the caregivers involved in the management of falls in the elderly that are targeted. More precisely, *MakerRehab©* will be addressed to physiotherapists, general practitioners and geriatricians, podiatrists, nurses, physical activity coaches and occupational therapists. In view of their workload, by offering a long and short assessment mode, with a design hypothetically conducive to naive users, *MakerRehab©* could fit well into their daily practice. By using the *UCDFmH*, the usability of *MakerRehab©* might be optimized. The innovative, open-source aspect of *MakerRehab©* could also help to promote access to

<u>Figure 13:</u> *Signal Processing Pipeline*



health care and delivery, by allowing technology to be transferred from the hospital setting to the home environment and outpatient practice and thus help to relieve hospital overcrowding. By adding full documentation and allowing raw data extraction, *MakerRehab©* could eventually stand out from similar technologies such as *Sway Medical Balance App©* for its trustworthiness.

In terms of Design and Technical issues: the fact that we are only dealing with a first prototype does not allow us to address this pillar in its entirety. However, the potential problems related to security, hardware, software and others, should be put in the forefront.

With regard to the Society: the constant search for patient consent could also answer the ethical issues that mHealth technologies are currently facing.

Limitations **——** This work is not without its limitations. First of all, the functioning of our predictive model raises questions about the multifactorial nature of the Full Assessment mode of *MakerRehab©*. Indeed, the feature importance analysis indicates that the model only considers the number of falls per year to predict the risk of falling and compute the balance score. Although crucial to the assessment of fall risk in clinical routine, future improvements of *MakerRehab©* should attempt to build a more balanced predictive model, where each of the 46 features would have a comparable importance score and thus a comparable contribution to fall risk assessment. As proposed by Fahimi et al (2021), future work should therefore focus on improving both the clinical relevance and the performance of the model[5]. Also, the context of the internship and the short time allocated did not allow us to involve the end-users in the design and development of *MakerRehab©*, although it is recommended by the *GDmH*. Future improvements should therefore include the use of surveys, questionnaires, and visits to health care facilities to constantly seek feedback from caregivers. This could greatly enhance the deployability of *MakerRehab©*.

Future work should also focus on the application of cyber security principles to ensure the protection of sensitive data such as health data used in *MakerRehab©*. This point is fundamental and inherent to the clinical translation of *MakerRehab©*.

Finally, the smartphone holder from *MakerRehab©* will also have to be improved in order to have a movable system that can be adapted to any size of smartphone.

CONCLUSION

Existing fall risk evaluation paradigms are either qualitative in nature (i.e., based on subjective conventional clinical tests) or lack of adoption (i.e., based on the current technologies) due to their poor transparency or accessibility. We here conceptualized, designed, and developed a novel faller identification system that could meet the challenge of current practices and technologies by combining conventional threshold-based methods with machine learning to screen (i.e., with the fall risk label) and monitor (i.e., with the balance score) falls-prone older adults, complementing conventional clinical assessment with an objectively measured functional test using the smartphone accelerometer.

Once the answers to the aforementioned limitations have been found, the next step includes validating the mobile application for clinical translation in elements such as intra- and inter-operator reliability, usability, or convenience.

When validated and further developed, *MakerRehab©'s* originality could not only improve the management of falls using a widely used technology, but also present an opportunity for unsupervised falls risk assessment.

Personal Note **——** This personal project allowed me to strengthen my knowledge in machine learning, project management, critical analysis of research articles, biomechanical data visualization and analysis, while learning the basics of clinical software development, user experience and interface (UI/UX) research, 3D printing and 3D modelling using Fusion360. While using Flutter for the first time, I also discovered a troubleshooting ability that I never knew I had before.

Although I need to continue working on my time management, these acquired skills will be central to my PhD and future career in the field of digital health technologies for rehabilitation.

Acknowledgements **——** I would like to thank my supervisor, *Kevin*, for his support and constructive criticism during my internship. In addition, I would like to thank the entire staff at *MakerLab* for their feedback and help throughout this unforgettable journey.

ANNEXE 1 : REFERENCES

1. Falls [Internet]. [cited 2022 Mar 19]. Available from: https://www.who.int/news-room/fact-sheets/detail/falls
2. Manchester D, Woollacott M, Zederbauer-Hylton N, Marin O. Visual, Vestibular and Somatosensory Contributions to Balance Control in the Older Adult. :10.
3. Nishiguchi S, Yamada M, Nagai K, Mori S, Kajiwara Y, Sonoda T, et al. Reliability and Validity of Gait Analysis by Android-Based Smartphone. Telemedicine and e-Health. 2012 May;18(4):292–6.
4. Amick RZ, Chaparro A, Patterson JA. Test-Retest Reliability of the sway Balance Mobile Application. JournalMTM. 2015 Jul 16;4(2):40–7.
5. Fahimi F, Taylor WR, Dietzel R, Armbrecht G, Singh Nb. Identifying Fallers Based on Functional Parameters: A Machine Learning Approach. In: 2021 IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE) [Internet]. Brisbane, Australia: IEEE; 2021 [cited 2023 Jan 5]. p. 1–6. Available from: https://ieeexplore.ieee.org/document/9718435/
6. Shumway-Cook A, Baldwin M, Polissar NL, Gruber W. Predicting the Probability for Falls in Community-Dwelling Older Adults. Physical Therapy. 1997 Aug 1;77(8):812–9.
7. Savadkoohi M, Oladunni T, Thompson LA. Deep neural networks for human's fall-risk prediction using force-plate time series signal. Expert Systems with Applications. 2021 Nov;182:115220.
8. Doheny EP, McGrath D, Greene BR, Walsh L, McKeown D, Cunningham C, et al. Displacement of centre of mass during quiet standing assessed using accelerometry in older fallers and non-fallers. In: 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society [Internet]. San Diego, CA: IEEE; 2012 [cited 2022 Apr 3]. p. 3300–3. Available from: http://ieeexplore.ieee.org/document/6346670/
9. Cerrito A, Bichsel L, Radlinger L, Schmid S. Reliability and validity of a smartphone-based application for the quantification of the sit-to-stand movement in healthy seniors. Gait & Posture. 2015 Feb;41(2):409–13.
10. Farao J, Malila B, Conrad N, Mutsvangwa T, Rangaka MX, Douglas TS. A user-centred design framework for mHealth. Schnall R, editor. PLoS ONE. 2020 Aug 19;15(8):e0237910.
11. Chatzipavlou IA, Christoforidou SA, Vlachopoulou M. A recommended guideline for the development of mHealth Apps. Mhealth. 2016 May 18;2:21.
12. Santos DA, Duarte M. A public data set of human balance evaluations. PeerJ. 2016 Nov 2;4:e2648.
13. Wollseifen T. Different methods of calculating body sway area. Pharmaceutical Programming. 2011 Dec;4(1–2):91–106.
14. Jin Huang, Ling CX. Using AUC and accuracy in evaluating learning algorithms. IEEE Trans Knowl Data Eng. 2005 Mar;17(3):299–310.
15. HAS. Chutes repetees personnes agees_-_argumentaire.pdf [Internet]. 2009 [cited 2023 Jan 23]. Available from: https://www.has-sante.fr/upload/docs/application/pdf/2009-06/chutes_repetees_personnes_agees_-_argumentaire.pdf

ANNEXE 2 : TIMELINE

Semester Project Cyrille

MakerLab

Cyrille MVOMO

Project Start:	lun, 9/19/2022
Project End:	fri, 2/03/2023

TASK	PROGRESS	START	END
September			
Set-Up | Flutter Practice | Literature Review	100%	9/19/2022	9/29/2022
October			
Flutter Practice | Literature Review | Pre-prototype	100%	10/3/22	10/31/2022
November			
Classifiers Building | Signal Processing |API building	100%	11/2/22	11/30/22
December			
App Development	100%	12/1/22	12/23/22
January/February			
App Development | Deliverables Preparation | Defense	100%	2/1/23	2/2/23

ANNEXE 3 : *MAKERREHAB©* DEMO

More Here : https://github.com/cyrillemvomo/FallRiskAPP