Fraunhofer
PORTUGAL

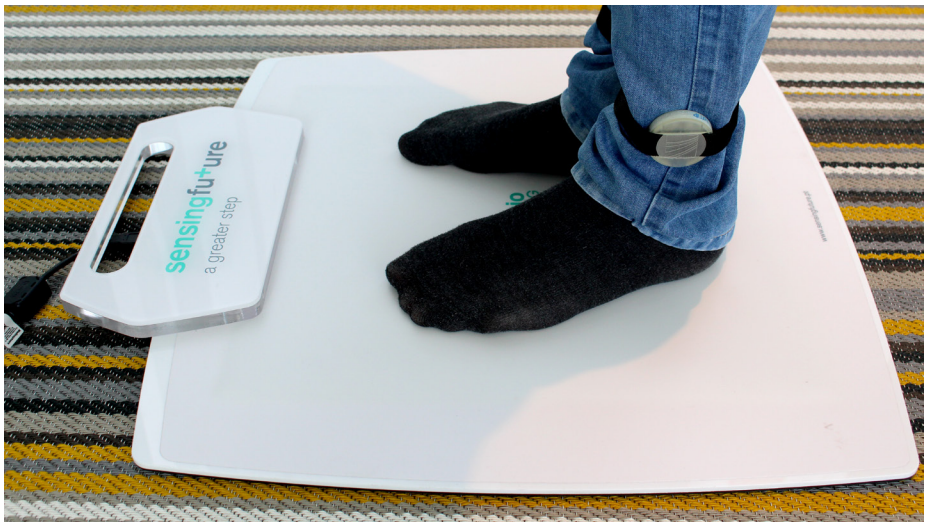

Fig1. Participant performing the screening instrumented tests.

FRSCREENING

FALL RISK SCREENING

The problem

Falls are a major problem that affects the elderly population, leading to institutionalization and dead. With physical, psychological and monetary consequences related, it is a field that has deserved a lot of attention and research.

More than action to diminish the consequences on the event of a fall, this problem requires prevention. To allow such prevention, the subjects that have higher risk for falls must be identified and exposed to a prevention plan.

Currently there are several tools and sets of tests that allow the classification of persons in high or low risk for falls groups. Some of these tools are based on physical tests, focusing on one or more risk factors for falls. There is no single tool that is currently accepted as the reference tool, that is easy to apply and that shows satisfactory results in all scenarios.

The solution

Fraunhofer Portugal in partnership with ESTeSC – Coimbra Health School, collected results on the execution of physical tests (that evaluated the presence of risk factors for falls), instrumented with inertial sensors and a pressure platform, from around 467 participants in several contexts. A follow-up period of six-months assured that information on post screening falls was also collected.

The proposed work involved the study of the applicability of Machine Learning classification algorithms in the process of identifying persons with high risk for falls.

Contact

Rua Alfredo Allen, 455
4200-135 Porto, Portugal

+351 220 430 300
info@fraunhofer.pt
www.fraunhofer.pt

Machine Learning

A Machine Learning system is a system that upon receiving some input information, processes its characteristics and reaches a conclusion. It is said that it learns by searching through a set of possible hypotheses to find the one that best represents the training information that is inputted to it.

Issues in Learning studied

- Imbalanced Classes: the dataset used for learning is imbalanced since it contains more information from one of the groups to classify than from the other. Several methods (balancing methods) were studied to overcome this issue
- Large number of features: the dataset has a large number of features for each one of the participants. This may increase computational effort while not improving results. Several methods (feature selection methods) were studied to overcome this issue

The existence of a database, potentiated with follow-up information on falls, allowed the application of such algorithms.

The objective was to build a model that showed good predicting capacities, understanding how Machine Learning could be incorporated in this process and which algorithms would perform better along with the study of how much the tests used for screening would contribute to the classification process.

Screening tests

The information collected from the participants, besides their personal information, included the performance on the execution of the Timed Up and Go test, the 30 Seconds Sit-to-Stand test, the 10 Meter Walk test, the Step test, the 4 Stage Balance test "modified", the X-Reach test and the Grip Strength test, instrumented with sensors.

From the 467 participants on the study, the follow-up information was known for 292, which information was used in this study. From these, 42 reported falls in the follow-up period.

There were 340 values of performance information (features) for each one of the participants.

Methodology

To understand the applicability of Machine Learning in this field, a set of tests that incorporated the widest range of algorithms and combination of methods was studied.

- The following classification algorithms (combined with balancing and feature selection methods) were applied: Decision Trees, Random Forests, K-Nearest Neighbors, Support Vector Machines, Naïve Bayes, AdaBoost, Neural Networks, Convolutional Neural Networks, RankSVM, RankBoost, RankNet, Cost Sensitive Random Forests and Easy Ensemble;
- The adopted methodology consisted on performing several tests with the combination of algorithms and methods, evaluating their performance for later comparison.

Each conducted test followed the steps presented in Fig2.

Results

The comparison between the several performances obtained revealed that the combination of Random Forests with the Near Miss balancing method using all values from the database except the ones from the execution of the Sit-to-Stand test led to the highest scoring performance, achieving 78% in the area under the receiver operating characteristic curve scoring measure.

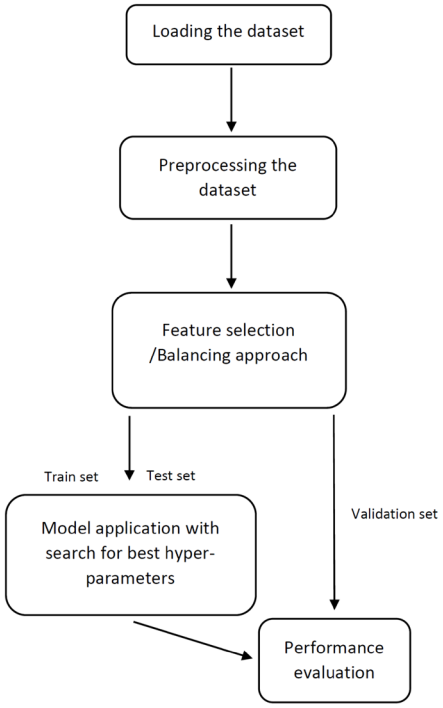

Fig2. Sequence of steps for each developed test.