M.SC.
MATHEMATICAL ENGINEERING
BIOMEDICAL ENGINEERING

POLITECNICO
MILANO 1863

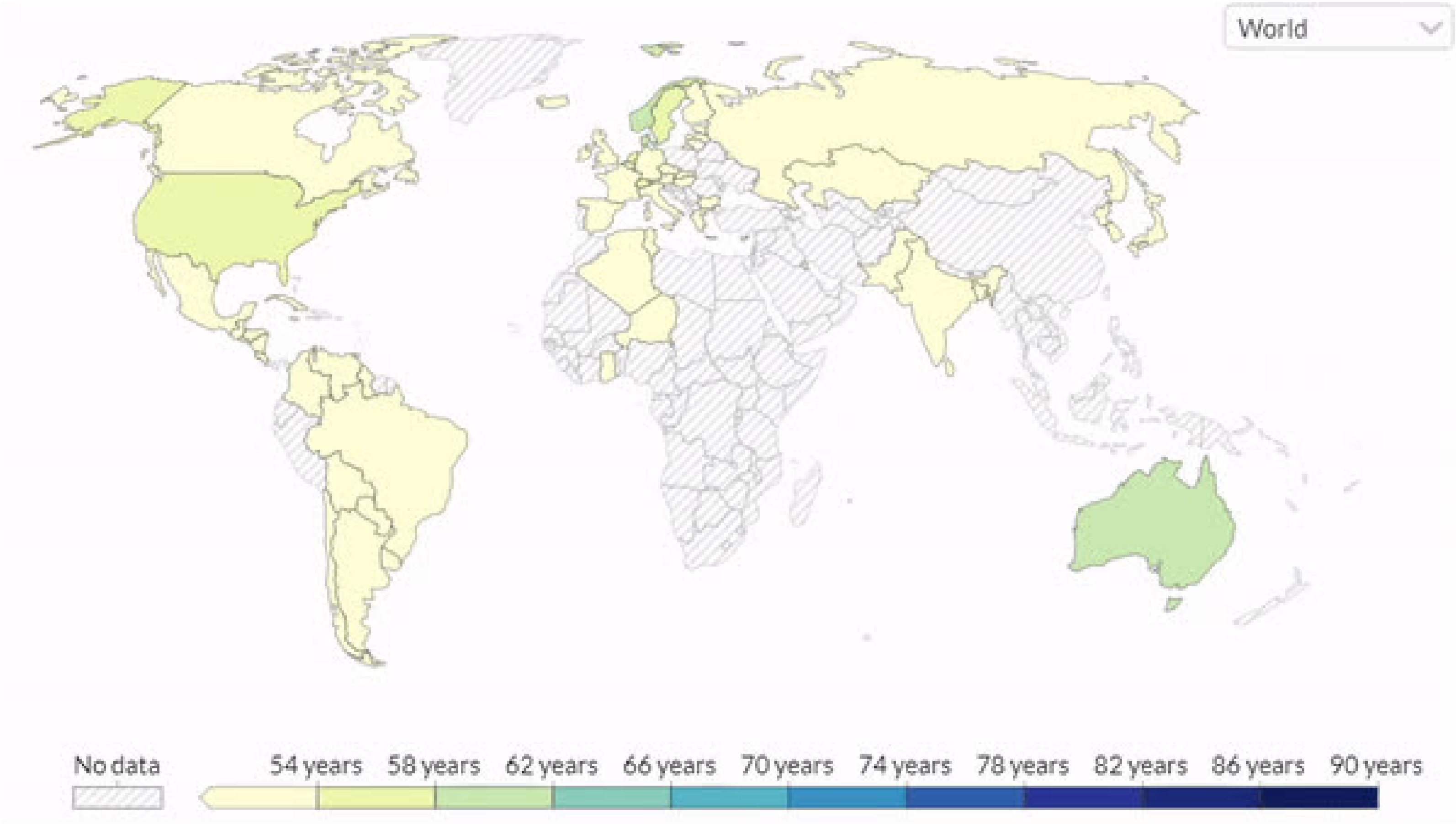
SLHD PROJECT CHALLENGE

CIPRIANI FILIPPO
PURICELLI ANDREA
SONNACCHI SOFIA

GOALS AND OBJECTIVES

Context

The global rise in life expectancy has posed the challenge of an increased risk of falls and of fall-related injuries among the elderly population. Having a high impact on healthcare systems, individuals, and society.
Wearable fall detection systems utilizing accelerometers and machine learning algorithms represent a promising solution to mitigate fall-related injuries.



Develop a machine learning algorithm able to identify patterns and distinguish between everyday activity and fall events using accelerometer data.

DATASET/CLASSES

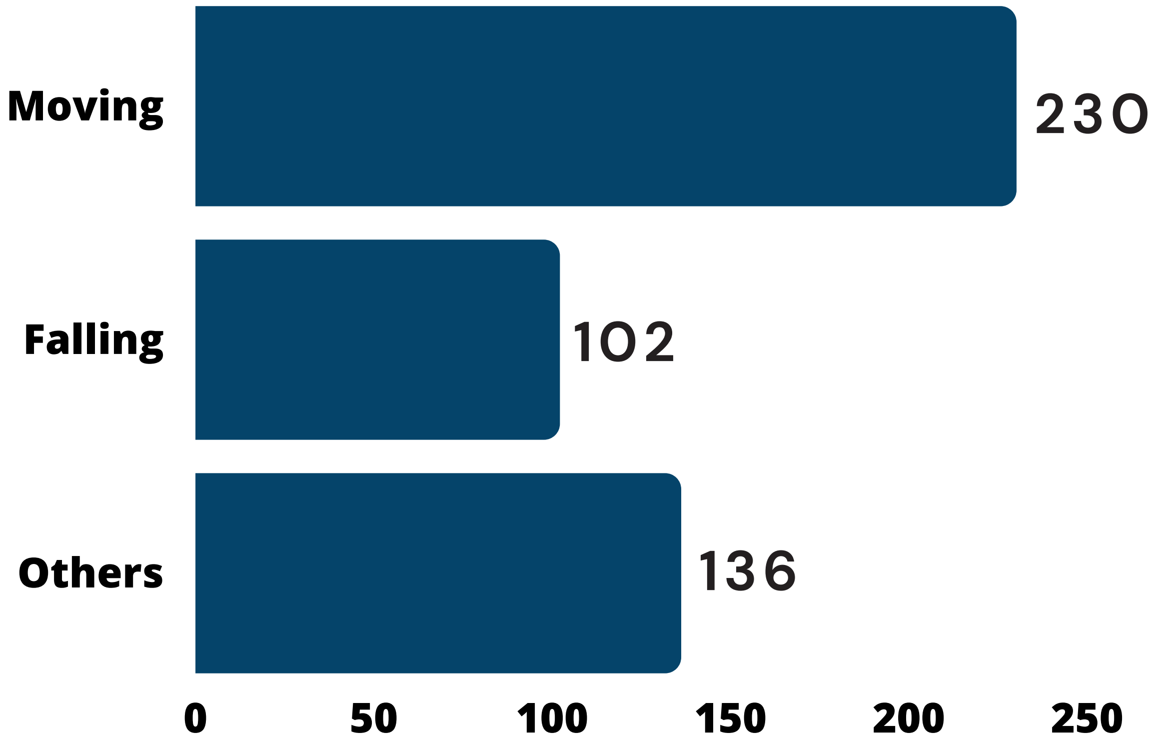
468 events:
- 12 types of activities of daily living (ADL)
- 3 types of simulated falls

Gathered from 16 individuals while wearing an inertial sensor unit attached sideways to their waist at a belt high. Each event includes the activity label and three-time series measurements corresponding to the accelerometer readings on the X, Y, and Z axis.

Moving
- Walking
- Running
- Jumping
- Down the stairs
- Up the stairs

Falling
- Forward fall
- Sideways fall
- Backward fall

Others
- Lying down
- Inactive



PREPROCESSING

MISSING VALUES

The dataset did not contain any missing values.

Before	After
104.47	0.0000
104.47	0.0050
104.48	0.0100

RESAMPLING

Consistent time interval for the accelerometer data. The signals were resampled at a frequency of 200 Hz and adjusted to begin at 0s

DISCARDED FDA

Due to the major challenges posed by the differing lengths of the signals, we opted for a traditional multivariate analysis, performing feature extraction on the time series.

DATA SCALING

Variables with different ranges were transformed to a common scale, allowing the confrontation of the features and their equal contribution in the analysis.

FEATURES EXTRACTED

60 features total

20 features extracted for each axis of the time series of the observation (X,Y,Z)

- Basic Statistical Measures (Max, Mean, etc.);
- Time Series Analysis inspired measures (Slope of linear fit, Zero Cross count, etc.);
- Signal Processing inspired measures (Area Under Curve, Fundamental Frequency, etc.).

PIPELINE

Models:

- Logistic Regression
- Random Forest
- SVM
- KNN
- Naive Bayes
- Ensemble

01 CORRELATION ANALYSIS

Removal of highly correlated features using Pearson correlation. Threshold: 0.7

02 MODEL FIT

Fit the model using the whole training set. Tune hyperparameters with GridSearch and evaluate performance with Cross-Validation

03 FEATURE SELECTION

Select the most important features using feature importances Threshold: elbow (FNR, Falling, Accuracy)

04 MODEL (RE-)FIT

Refit the model with selected features. Evaluate on test set

RESULTS

- Logistic Regression
- Random Forest
- Ensemble Model (LR+RF, soft voting)

Model	Precision	Recall	Features
Logistic Regression	0.88	0.94	13
Random Forest	0.86	0.97	4
Ensemble	0.94	0.97	16

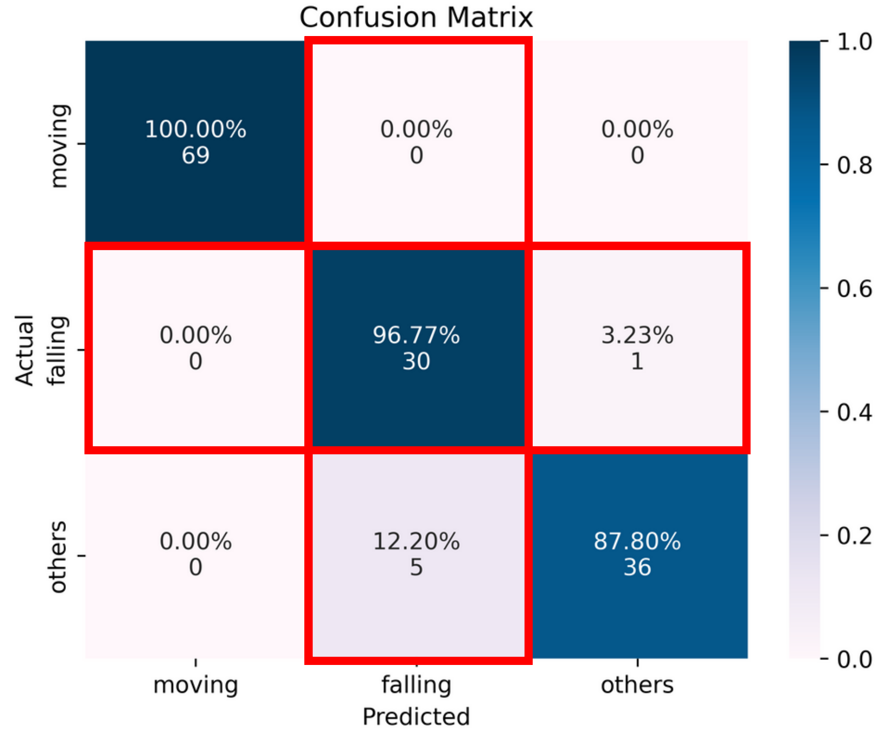
96%

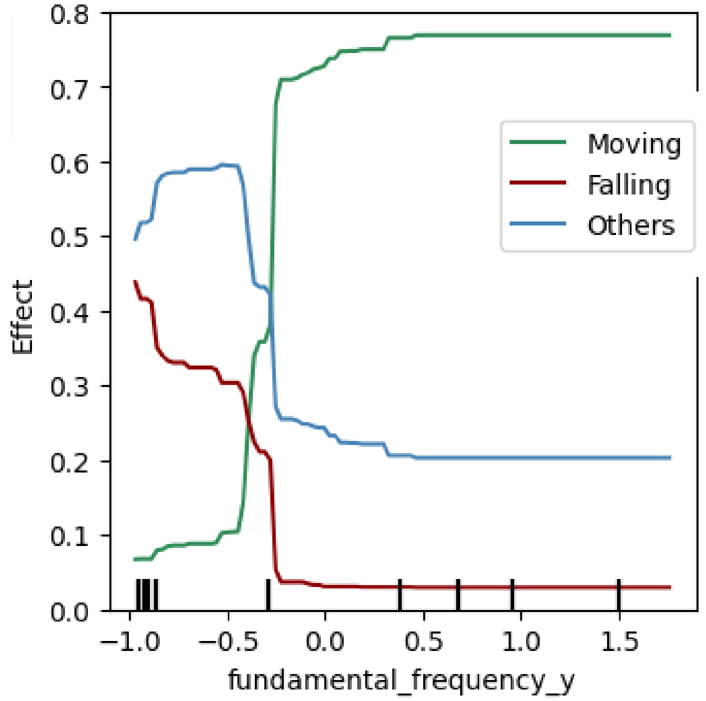
ACCURACY

FINAL MODEL: RF

4 FEATURES

- Fundamental Frequency (Y-axis)
- Mean (Y-axis)
- Minimum (Y-axis)
- Kurtosis (X-axis)

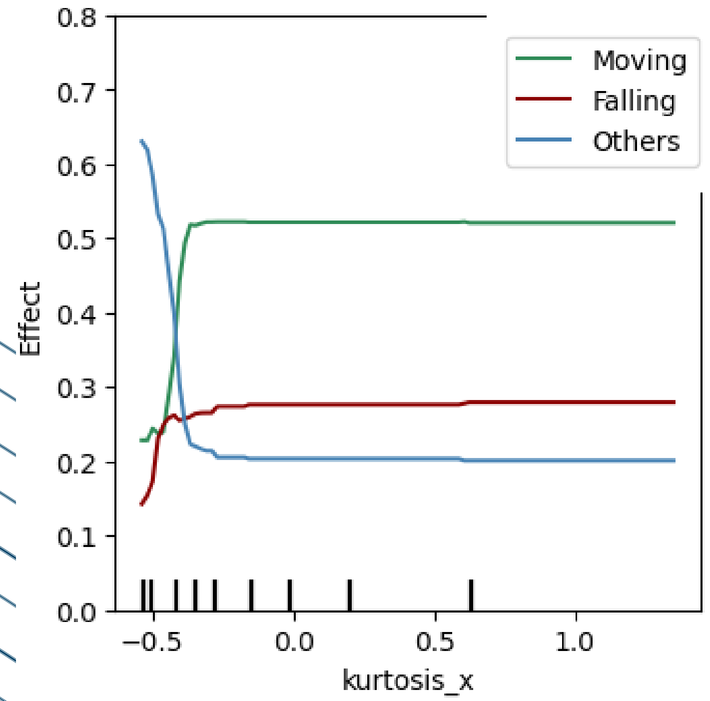


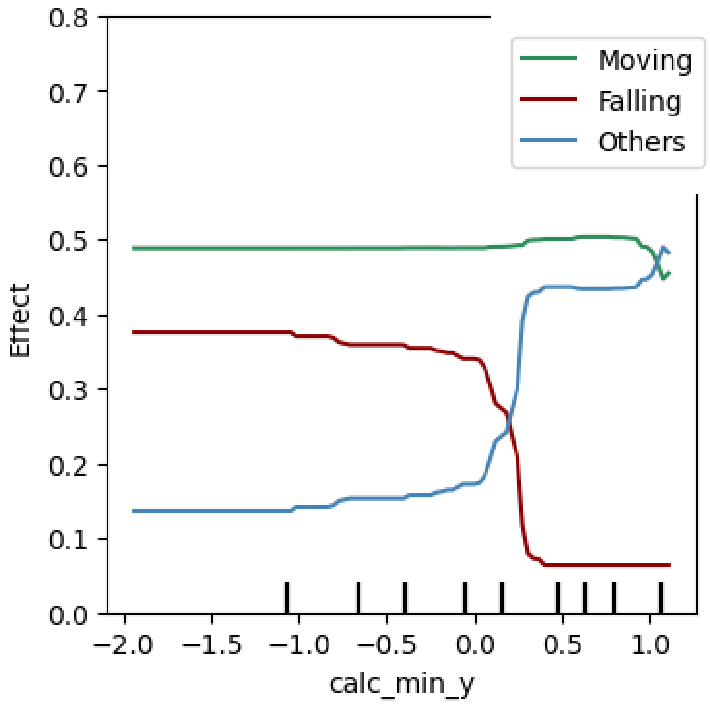


INTERPRETATION

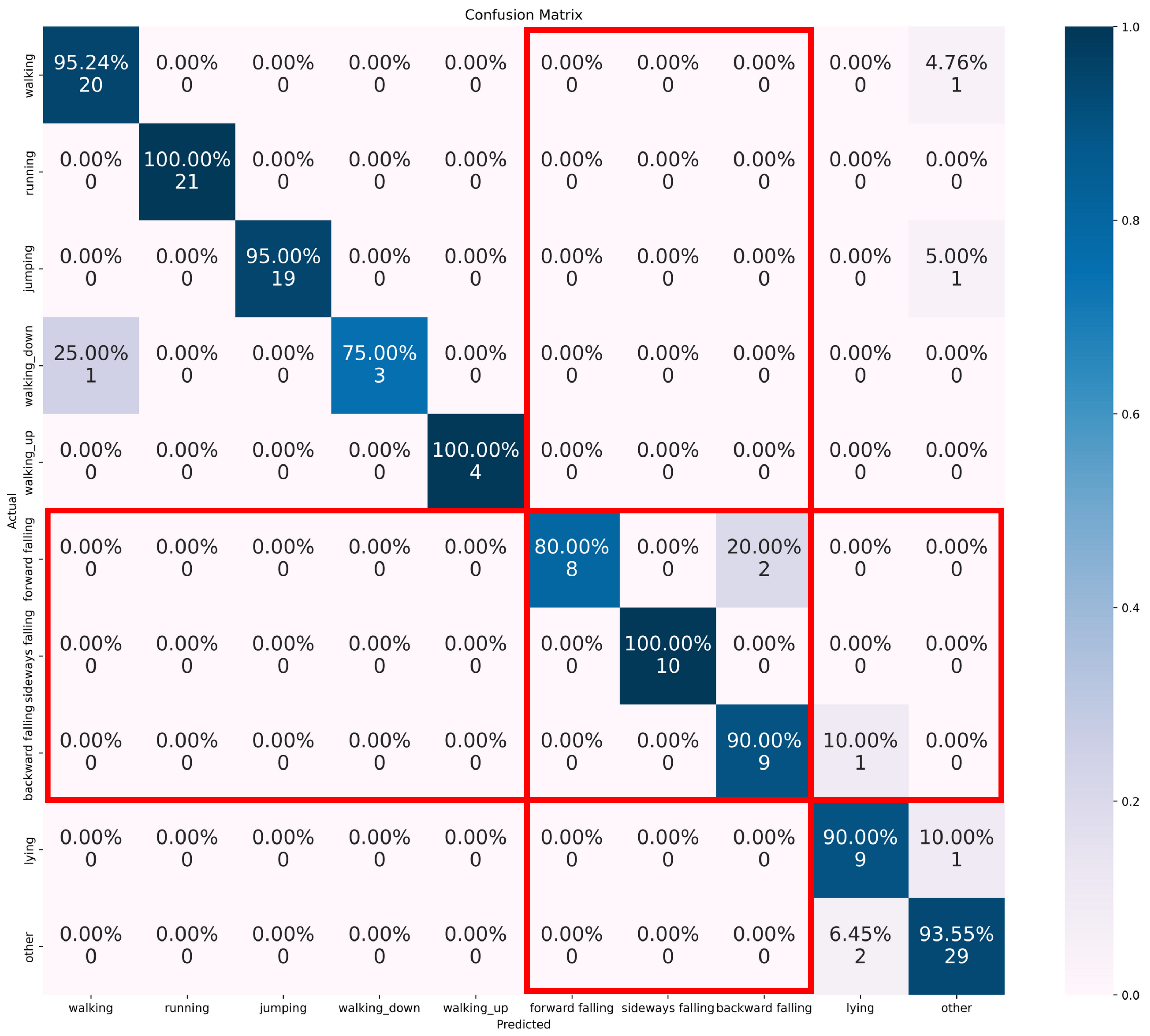
Using partial dependence plots:

- Fundamental Frequency (Y-axis) separates well the 3 categories
- Kurtosis (X-axis) divides Moving from the other two classes
- Minimum (Y-axis) splits Falling from the other two classes





10-CLASSES: ENSEMBLE (RF + LR)



14 FEATURES

- 6 Features for the X-axis
- 6 Features for the Y-axis
- 2 Features for the Z-axis
- Fundamental Frequency and Minimum (Y-axis) are still among the most influential features

Model performances

Model	**Class**	Precision	Recall	F-score	CV-Acc	**Features**
	Walking	0.95	0.95	0.95		
	Running	1.00	1.00	1.00		
	Jumping	1.00	0.95	0.97		
	Walking down	1.00	0.75	0.86		
Ensemble	Walking up	1.00	1.00	1.00	92.74%	14
	Forward fall	0.89	0.80	0.84		
	Sideways fall	1.00	1.00	1.00		
	Backward fall	0.82	0.90	0.86		
	Lying down	0.90	0.90	0.90		
	Inactive	0.91	0.97	0.94		

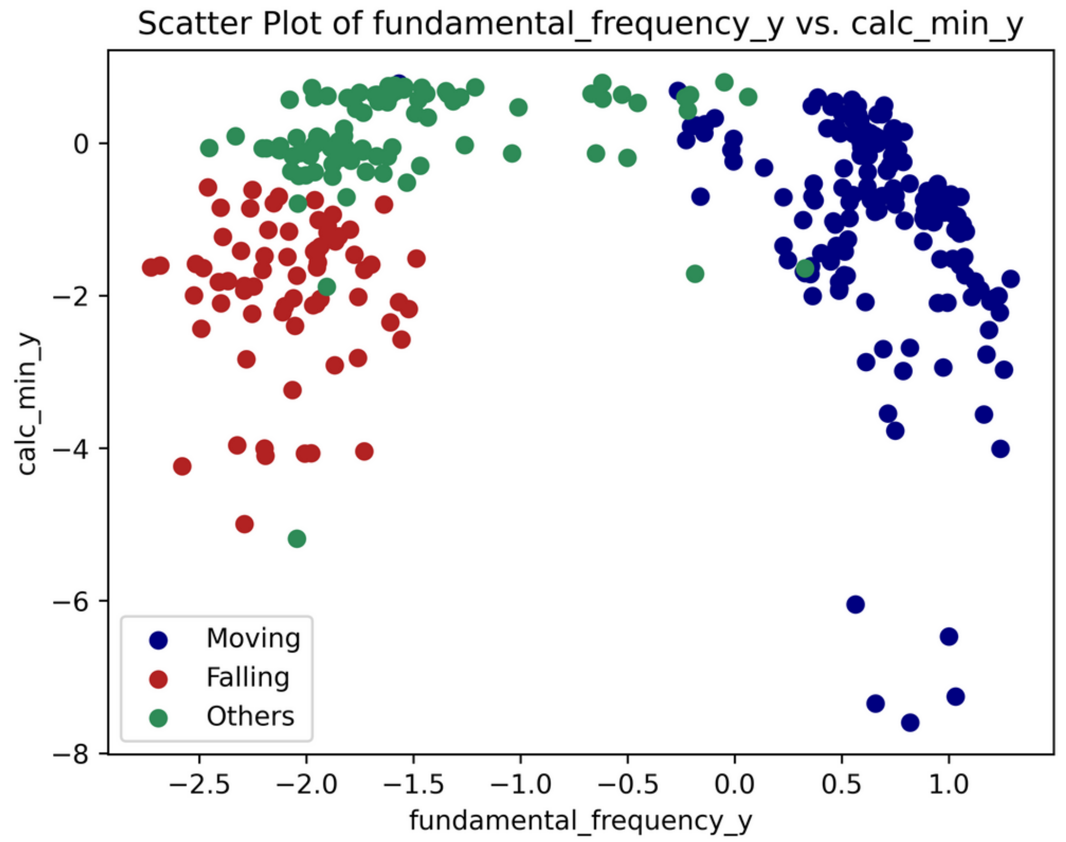
CONCLUSIONS

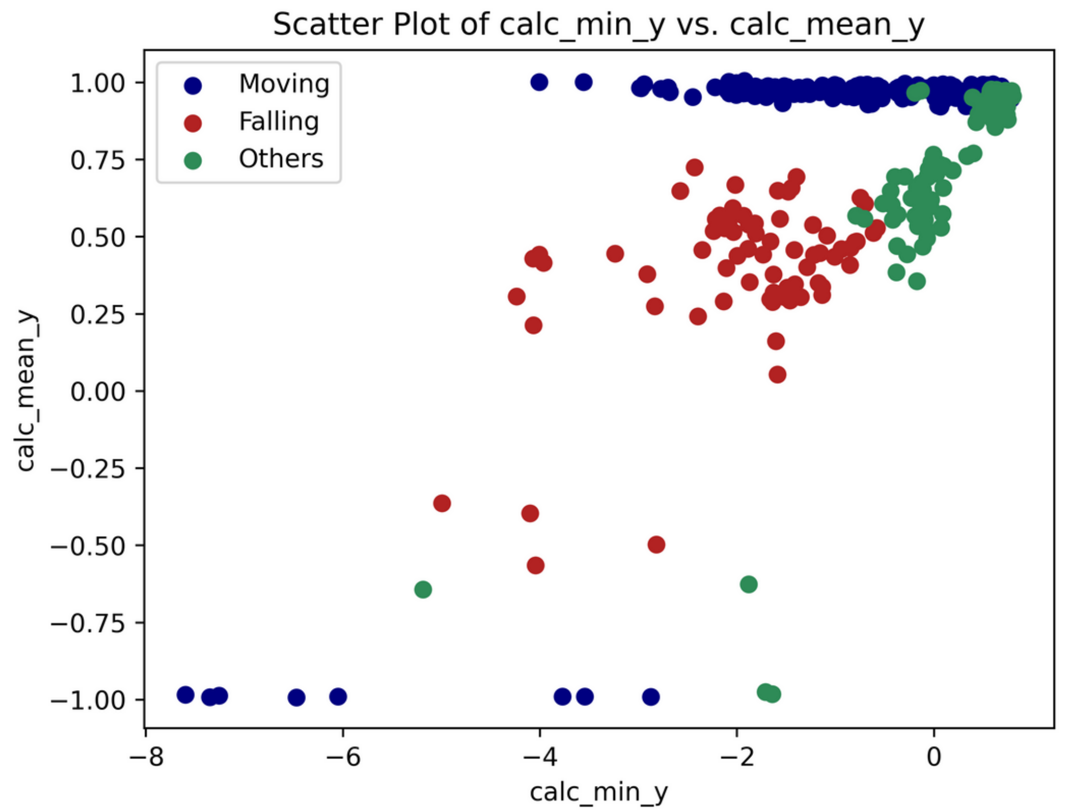
MODELS

We tried many models and we achieved an overall good performance, especially with the Random Forests.

STRENGTH OF THE MODEL

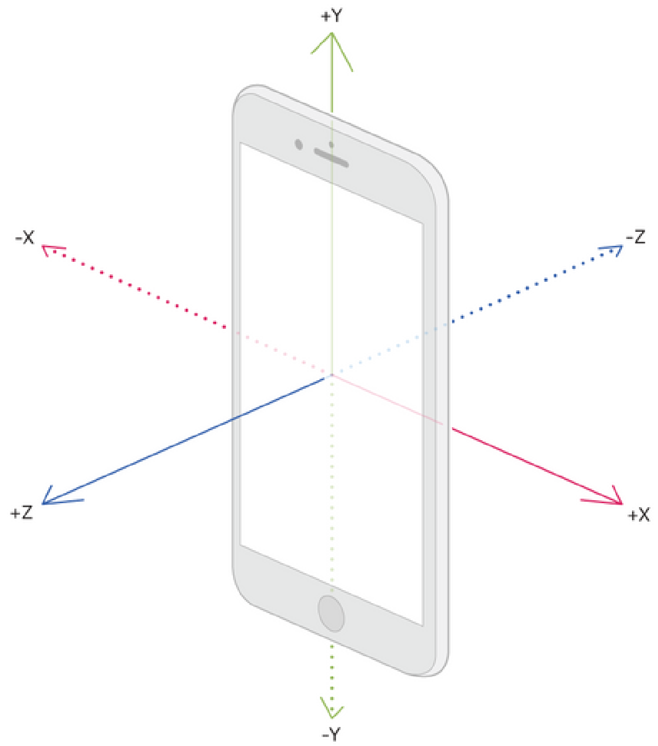
We found meaningful features such as the fundamental frequency that allowed us to cluster the 3 classes accurately.





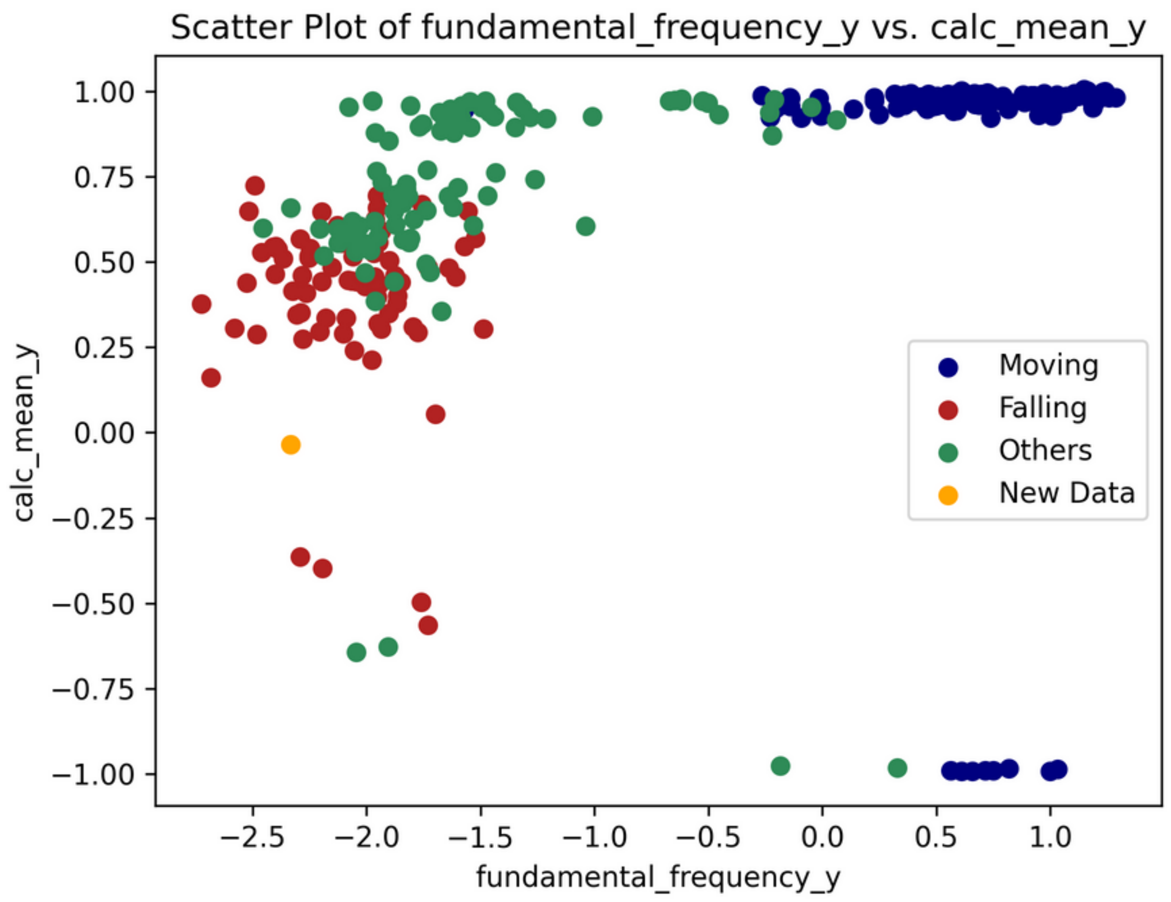
TESTING THE MODEL
(AGAIN)

Using our phones' accelerometer we collected new data.



Activities:

- Walking (up-down the stairs, running)
- Falling (simulated sideways, phone drop)
- Lying



IMPORTANT STEPS:

- Resampling to match frequency
- Scaling

THANKS FOR LISTENING

GitHub

REFERENCES

1. url: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health#:~:text=By%5C%202030%5C%2C%5C%201%5C%20in%5C%206,will%5C%20double%5C%20(2.1%5C%20billion.
2. url: https://www.who.int/news-room/fact-sheets/detail/falls.
3. Curtis S Florence et al. 'Medical costs of fatal and nonfatal falls in older adults'. en. In: J. Am. Geriatr. Soc. 66.4 (Apr. 2018), pp. 693–698.
4. Raul Igual, Carlos Medrano and Inmaculada Plaza. 'Challenges, issues and trends in fall detection systems'.en. In: Biomed. Eng. Online 12.1 (July 2013), p. 66.