Proof of Concept for the Development of Digital Biomarker using Raw Accelerometer Data from a Wrist Worn Device

Louis Smith, BAI, MSc; Marie McCarthy, MSc, MBA; Michael Phillips, PhD; Willie Muehlhausen, DVM. ICON plc

Objective
We wanted to conduct a proof of concept for developing digital biomarkers from raw accelerometer data. Accelerometers have the ability to capture significant quantities of raw data, potentially containing patterns which, if discoverable, could be used to quantify specific motor movements. The ability to detect such movements has value in identifying periods of tremor or other neuromuscular disorder.

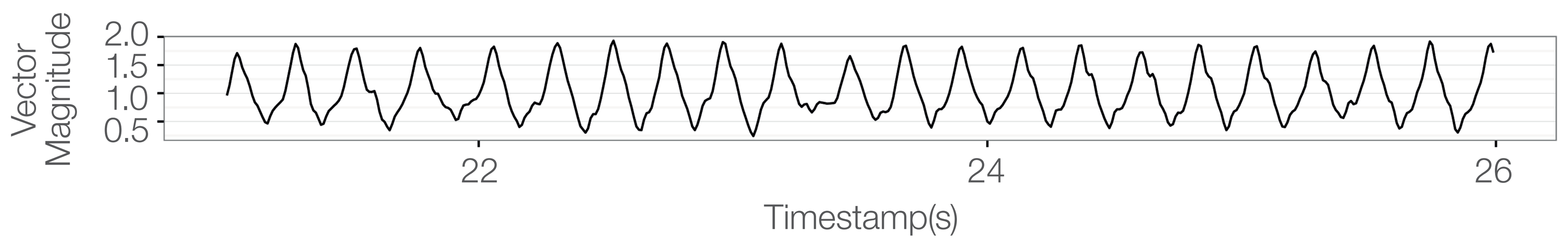
Method
Two subjects, A and B, each wore an accelerometer device on their dominant wrist for 72 and 18 hours, respectively. They performed multiple teeth brushing events and kept an event diary noting the times of brushing.

Vector magnitude was generated from the X, Y and Z raw 100 Hz accelerometer data, and a brushing label "yes" or "no" was applied by at minute level.

The raw data was reduced into 60 second summary epochs. The summary data was centred on the vector magnitude statistics and pattern descriptive metrics.

See plot below of raw teeth brushing pattern which we want to discover and an example of summarised data for a single epoch.

5 second Example of Vector Magnitude Pattern for Teeth Brushing



5 second sample of raw 100 Hz teeth brushing pattern.

Example of summary data for a single epoch

Subject	Epoch_ ID	Min	Qu1	Median	Mean	Qu3	Max	Peaks /sec	Mean Period	SD Period	Mean Amp	SD Amp	Brushing
A	20/01/2016 11:46:00	0.220	1.006	1.013	1.019	1.018	4.558	29.78	0.033	0.030	0.019	0.095	no

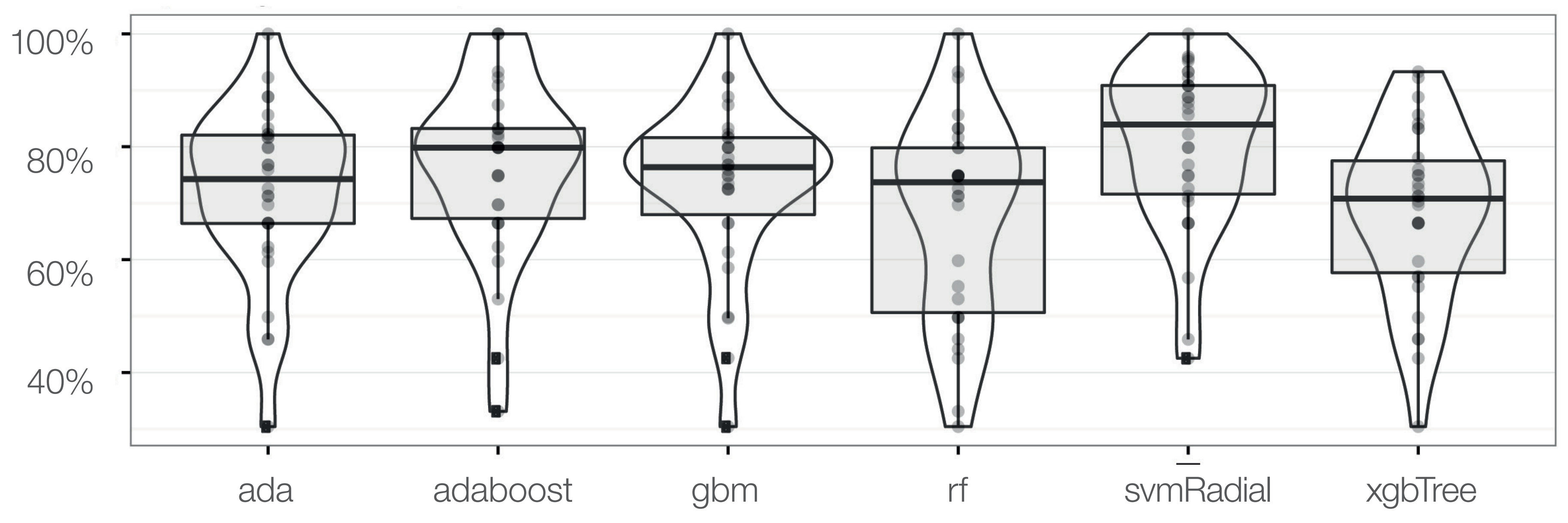
The summary data for subject A was partitioned into training (for building models) and testing (for evaluating models) datasets with a 60/40 split; data from subject B was withheld from the model to be used later to investigate the potential for generalizing the model to unseen subjects.

A set of classical machine models (decision trees, random forest, etc.) were built on the training dataset.

Results
Due to the limited data, unbalanced nature of the classes and the length of the epoch, the performance of any model will be dependent on the training/testing split. Over 30 different splits of training/testing data, the Support Vector Machine (svmRadial) model showed the highest median Kappa, with a distribution skewed towards 100%.

Kappa Results from Models-30 Data Partitions

The evaluation results of models trained & tested on Subject A shows significant variation in performance depending on the data split.



Distribution of model kappa results for 30 splits of data – subject A

Confusion Matrix and Statistics for svmRadial Model

Reference						
Prediction	n	y	Accuracy:	0.9981	Sensitivity:	0.9991
n	1068	1	95%CI:	(0.9933, 0.9998)	Specificity:	0.8889
y	1	8	Kappa:	0.888	Pos Pred Value:	0.9991
					Neg Pred Value:	0.8889

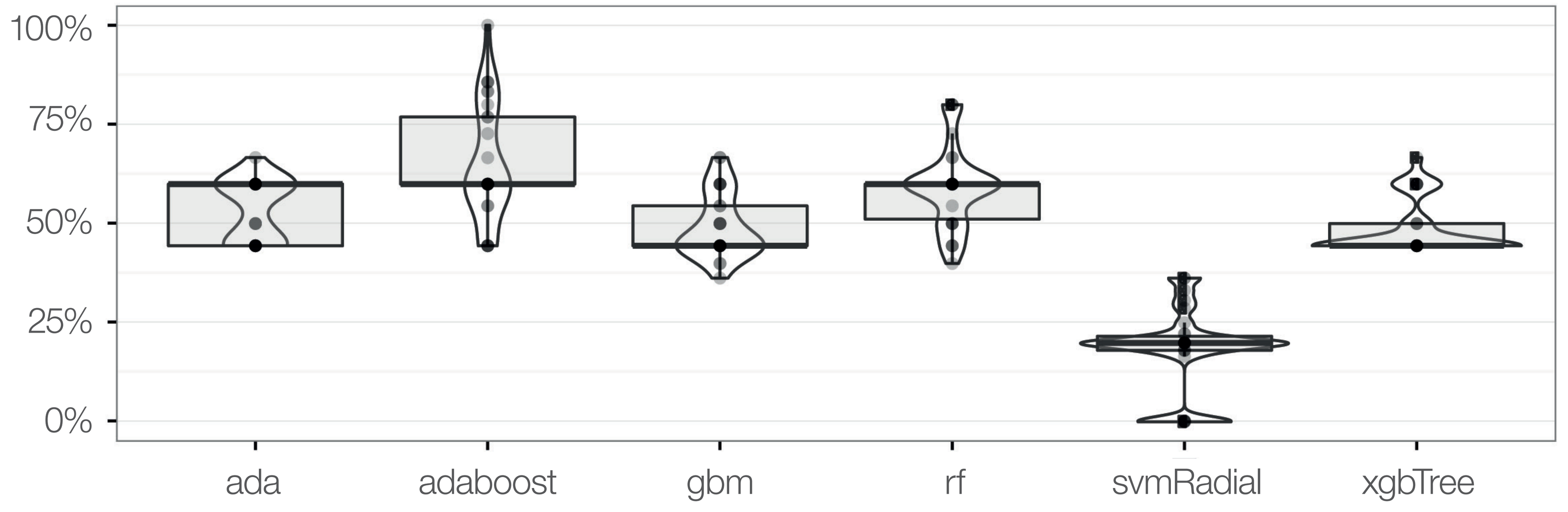
Sample results of an svmRadial Model on subject A test data

In order to investigate if the models had the potential to recognise similar patterns in unseen subjects, the same 30 models were used to classify the epoch summary data from subject B.

The two subjects were very different in terms of physical attributes and different teeth brushing patterns would be expected. In this instance, the predictive power of the svmRadial model was the weakest; however, some predictive power remained, particularly for the Adaptive Boosting (adaboost) model.

Kappa Results for Predicting Unseen Subject-30 Data Partitions

Using model trained on Subject A's teeth pattern and using it classify periods of brushing for Subject B shows some predictive power. Variance shows sensitivity to the training/testing partition



Distribution of model kappa results for 30 splits of data – subject B
Note: change of scale on Y-axis

Confusion Matrix and Statistics for adaboost Model

Reference						
Prediction	n	y	Accuracy:	0.9979	Sensitivity:	0.9986
n	1392	1	95%CI:	(0.9938, 0.9996)	Specificity:	0.8333
y	2	5	Kappa:	0.7682	Pos Pred Value:	0.9993
					Neg Pred Value:	0.7143

Sample results of an adaboost Model on subject B data

Conclusions
This POC has demonstrated that it is possible to use machine learning techniques to train a classification model from summarized raw accelerometer data to identify periods of specific movement patterns. The quantity of data required to build a robust algorithm will depend on the variance in the pattern between individual events and from subject to subject.

A model which performs best on the evaluation dataset, may not necessarily maintain that performance level on unseen subjects.This approach has potential application in objectively measuring motor movement events in neuromuscular disorders but also in the development of unique personal digital fingerprints.

Disclosure — The authors have declared there are no conflicts of interest.