Nestlé — Good food, Good life

Predictors Optimization for Sensory Profiles Modelling Based on Electronic Signals

Jean-Vincent Le Bé – Nestlé Research

Salzburg, July 9th 2024

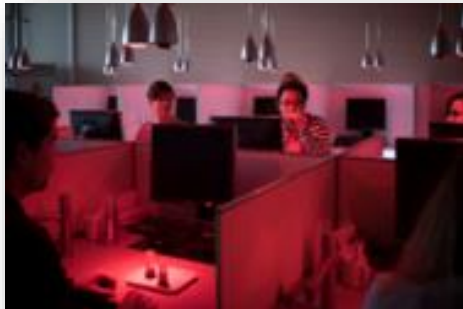# E-tongue enables similarity / dissimilarity identification of wider range of products

Sensory panel limitations
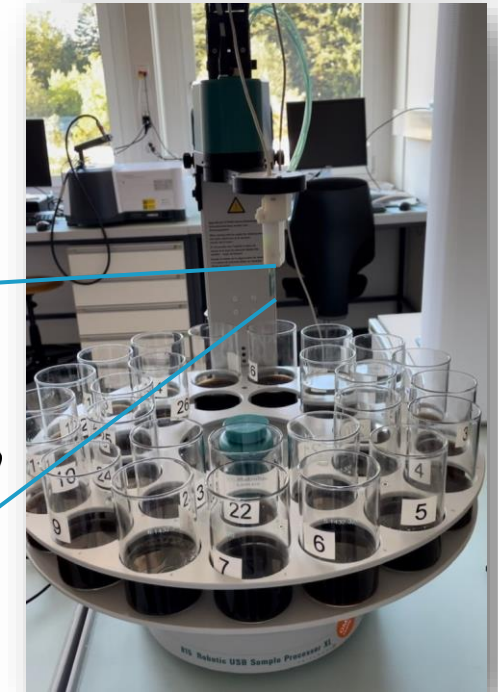- ❖ Training
- ❖ Availability
- ❖ Fatigue
- ❖ Food safe product

E-tongue
- ✓ More samples in a session
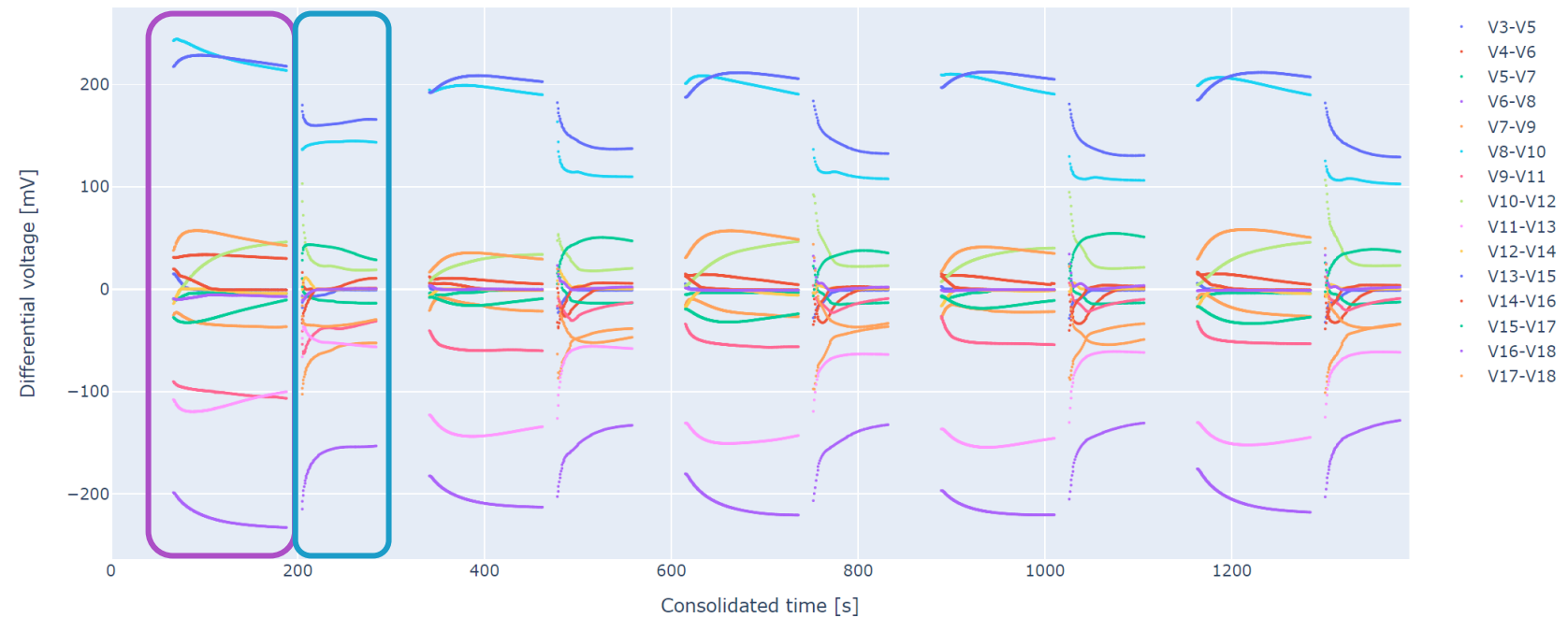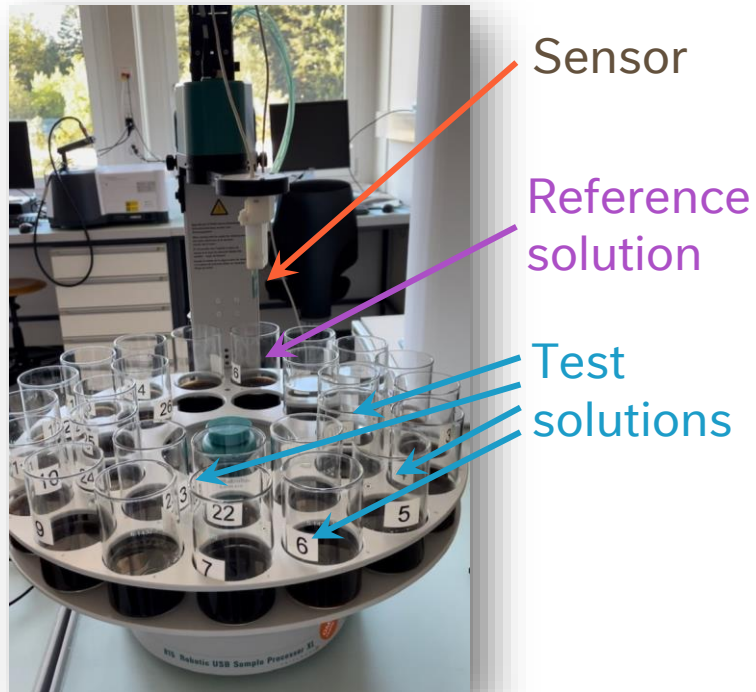- ✓ No need for food grade product
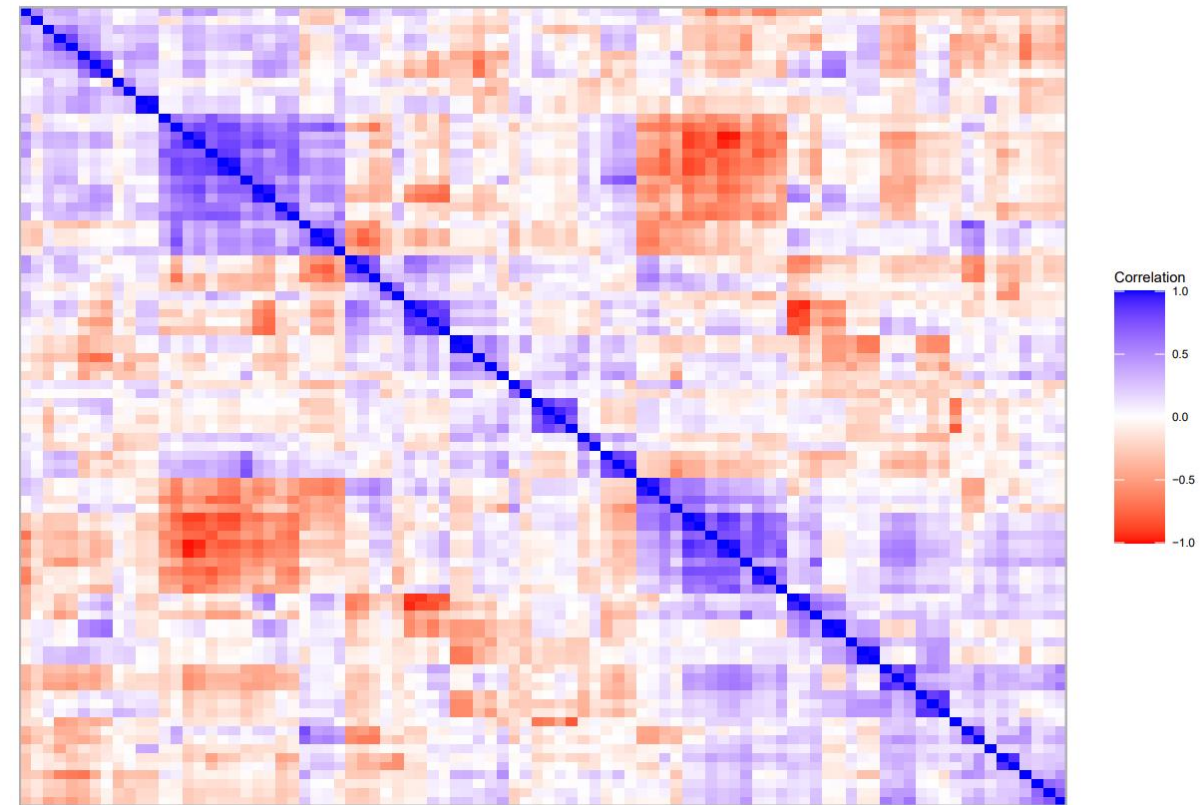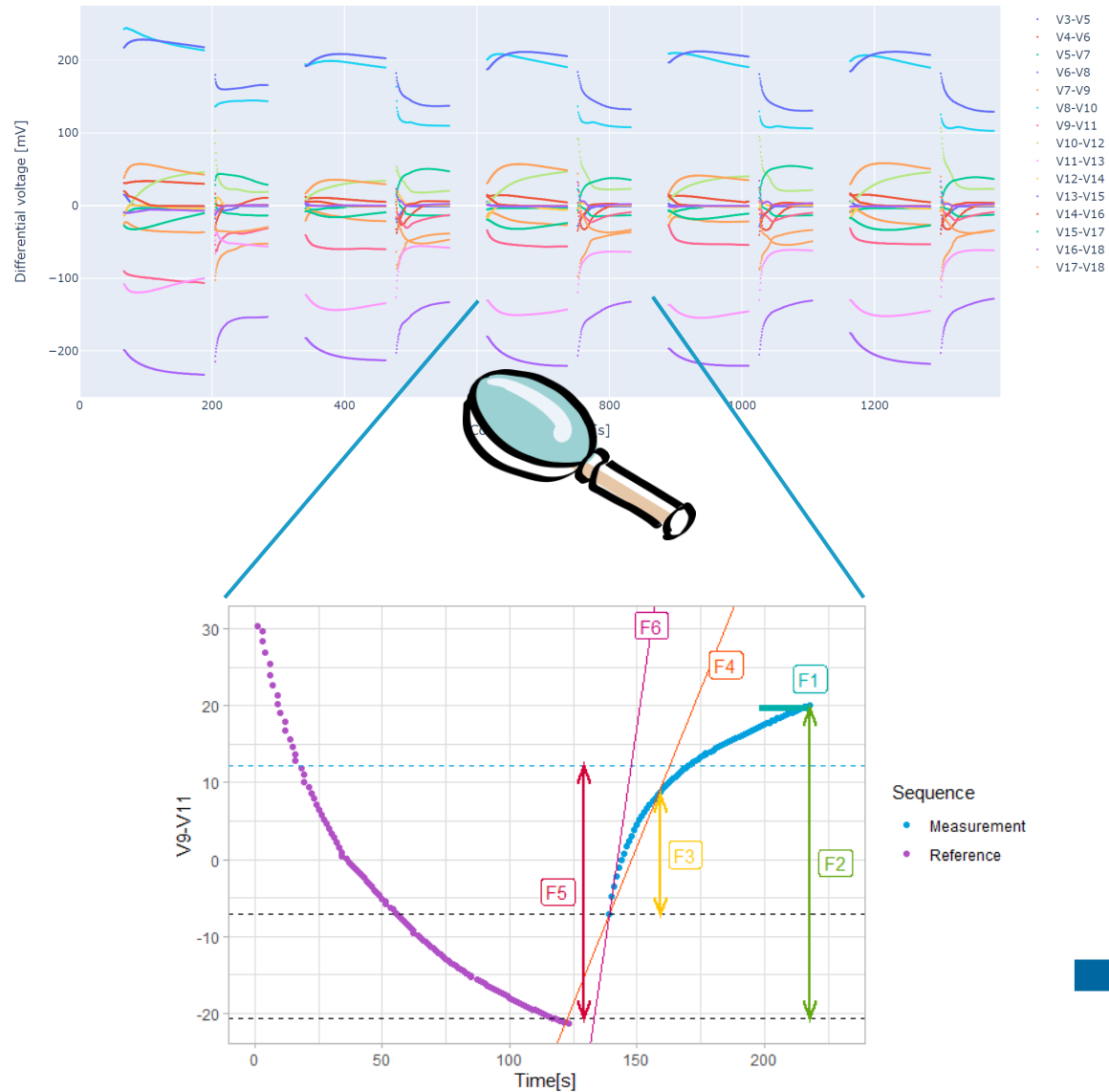- ✓ Close to continuous operation

Electrode array

BUT
Human sensory system is complex

Research and Development

16 electrodes coated with selective bio-polymers as proxy to physiological sensation ➡ 15 voltage signals

Research and Development

15 time series each transformed into 6 defined features

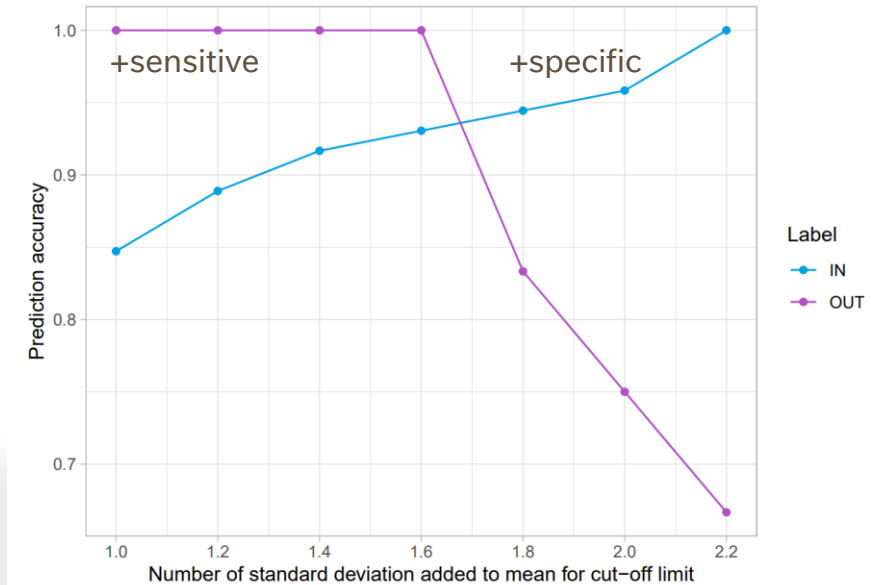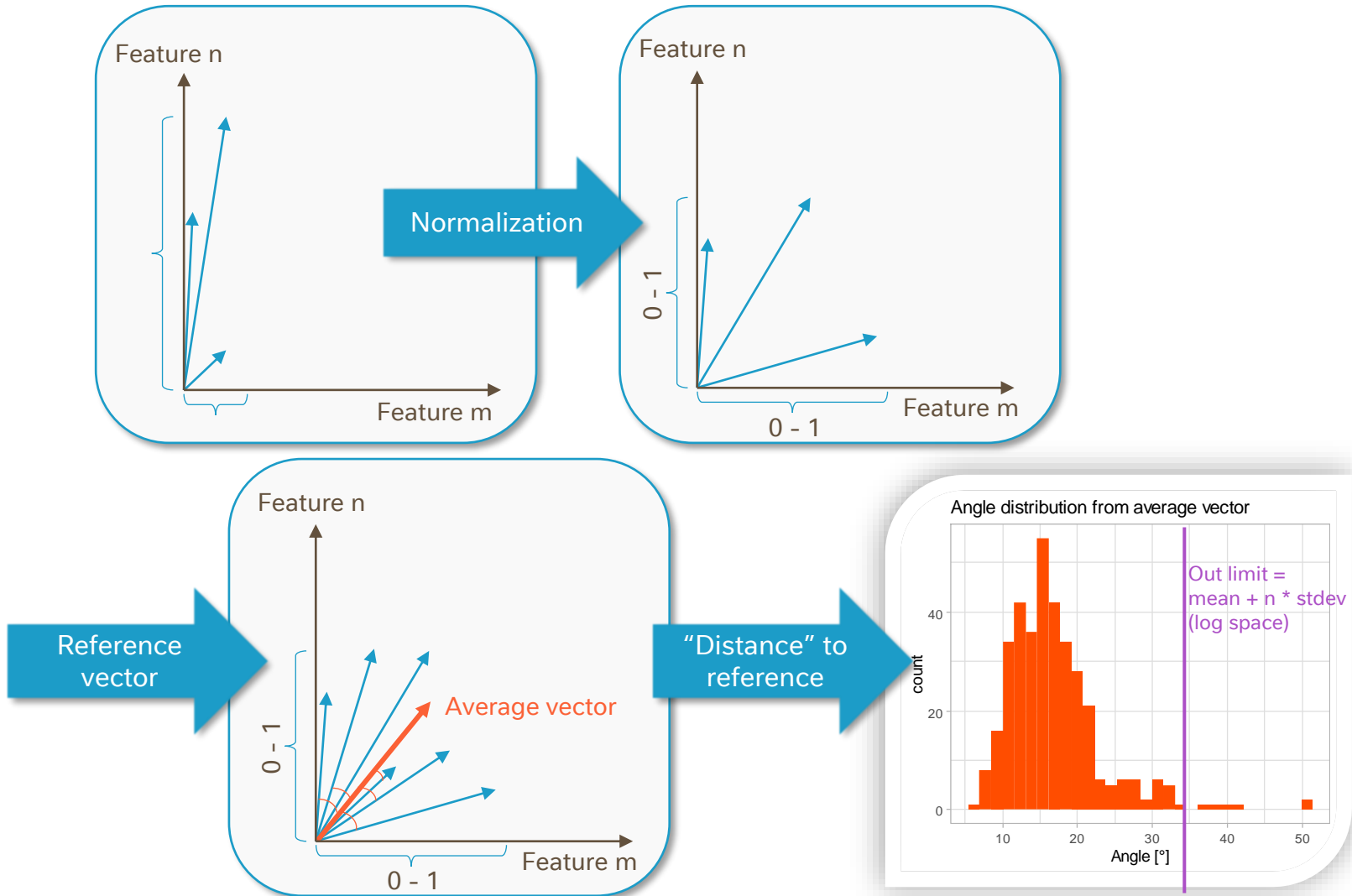➡️ **90 variables with many correlations**

Research and Development

Random Forests variable importance shows number of features can be reduced

❖ Gini index distribution on 100 random forests

❖ 4 batches recorded: model trained on 3 (leave one out)

❖ Arbitrary choice of a cut-off limit at 1 in at least one batch

❖ F1, F2 and F5 for each of 15 series sufficient
(blue vs black box plots)

Variable Importance by 100 Random Forests

Features
— F1 | F2 | F5 only
— Others

Features Batch 1
Features Batch 2
Features Batch 3
Features Batch 4

Research and Development
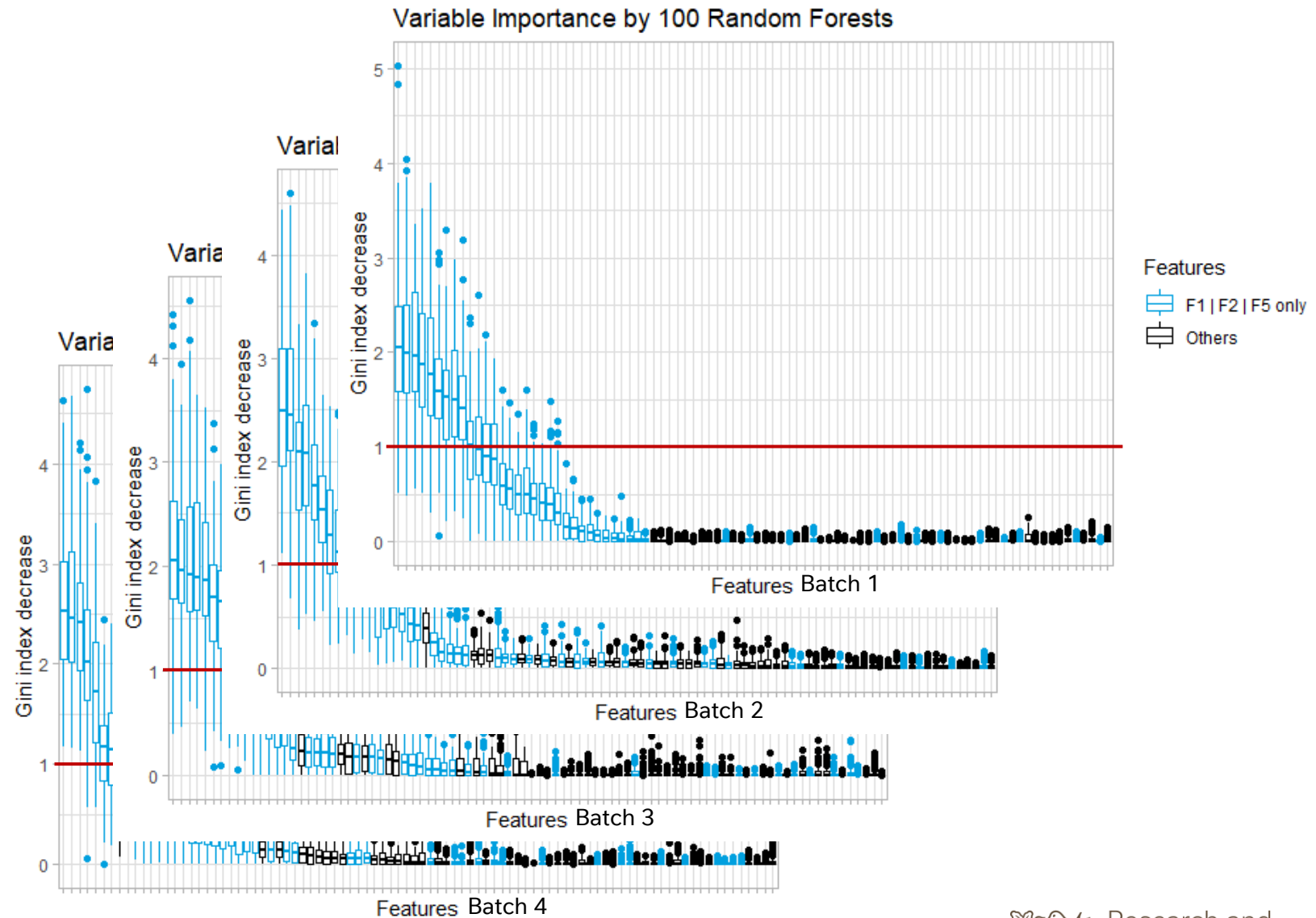
Angle Classifier can be trained on a single category
Is product under test IN or OUT of the category?

Feature n → Normalization → Feature n (0 – 1) (0 – 1) Feature m

Reference vector → Feature n, Average vector, 0 – 1, Feature m (0 – 1) → "Distance" to reference

Angle distribution from average vector

Out limit = mean + n * stdev (log space)

count: 40, 20, 0

Angle [°]: 10 20 30 40 50

+sensitive   +specific

Prediction accuracy: 1.0, 0.9, 0.8, 0.7

Number of standard deviation added to mean for cut−off limit: 1.0 1.2 1.4 1.6 1.8 2.0 2.2

Label
— IN
— OUT

Positive: Labelled as OUT
Sensitivity > Specificity

Research and Development

Dimension reduction and angle classifier enables good prediction accuracy with only data from a reference population
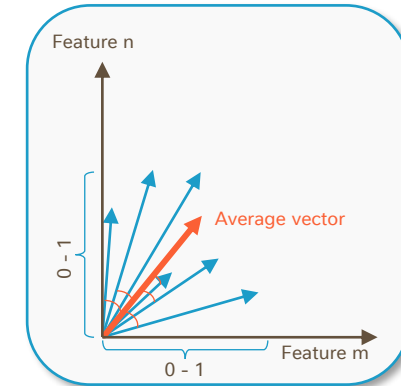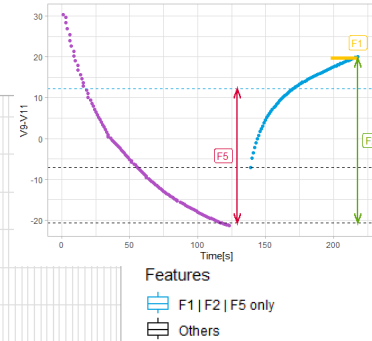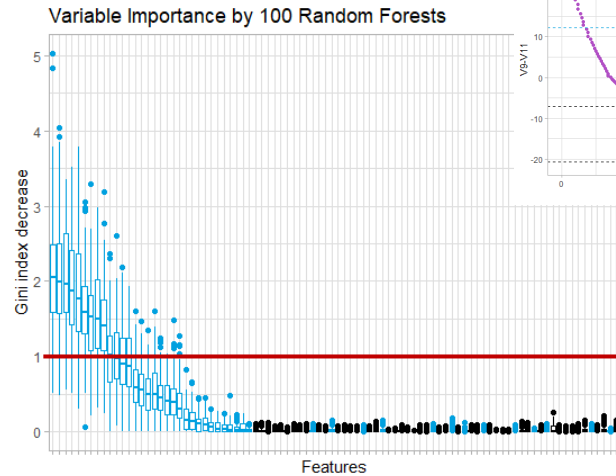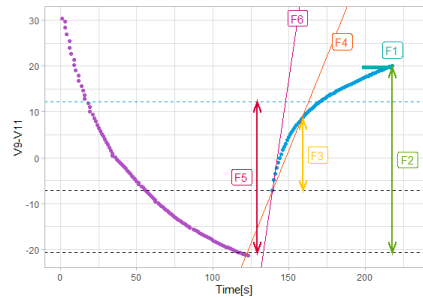
100% sensitivity with
> 90% specificity

Research and Development