

Activity recognition by sweat intensity

Maxim Grishin
@maximgrishin

THE HARDWARE

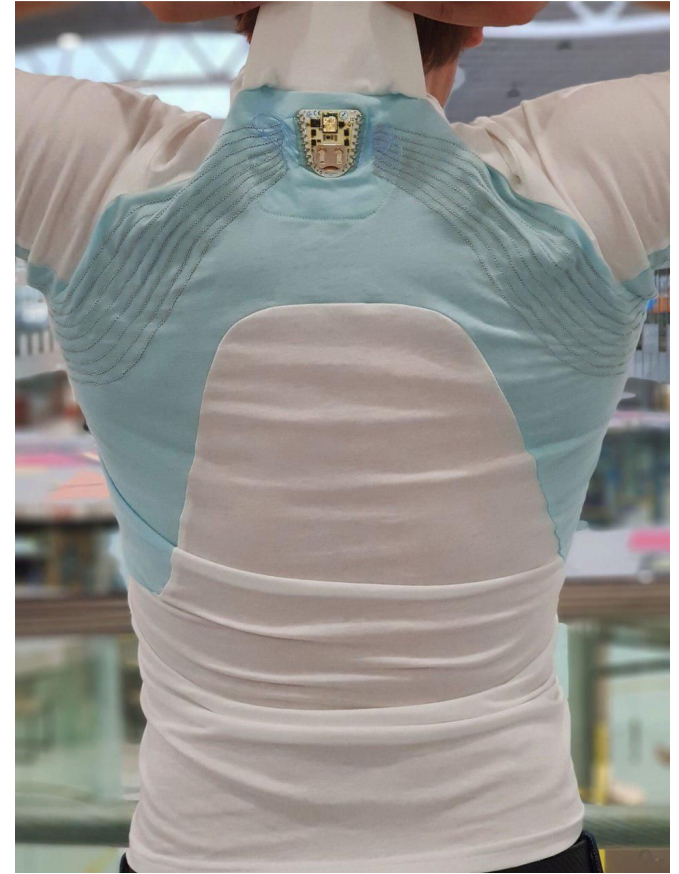


**REAL
LIFE
TESTED**

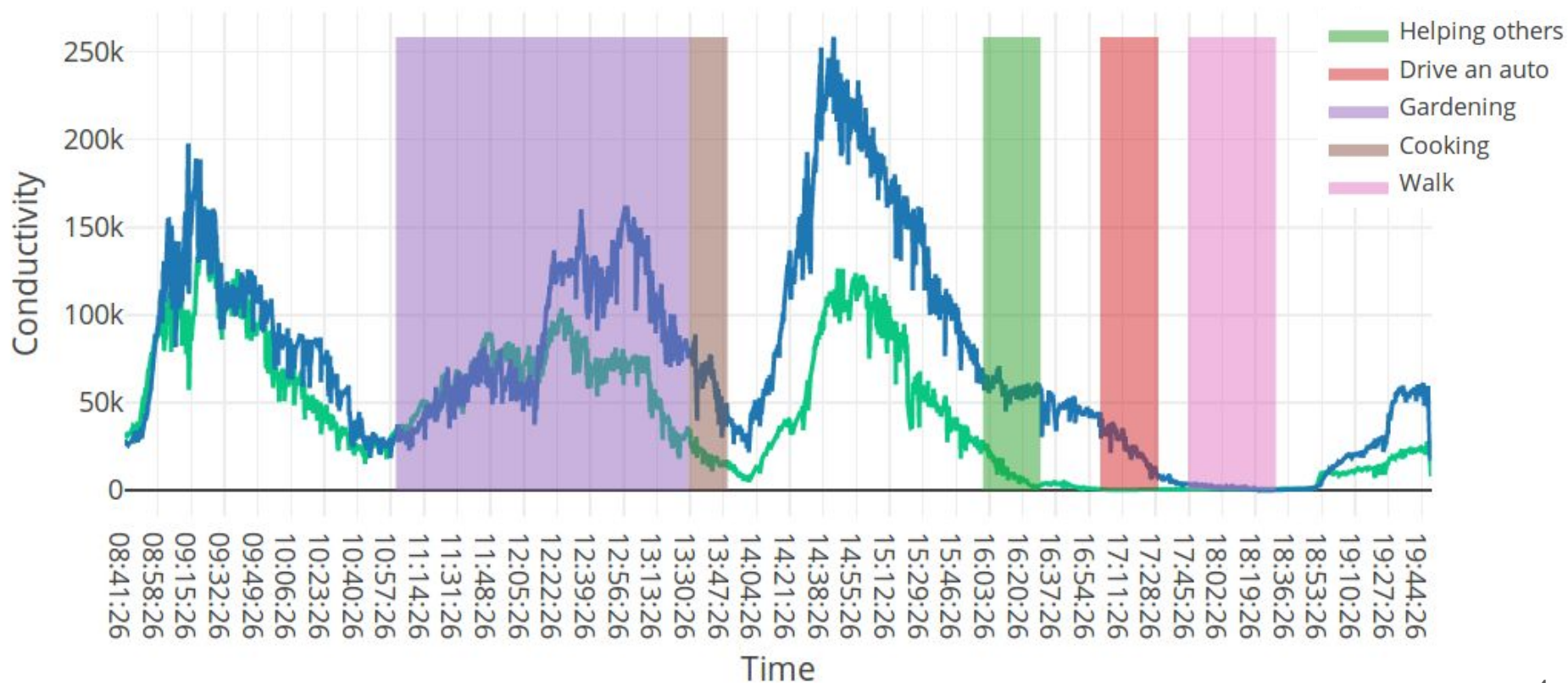


SMART SHIRT

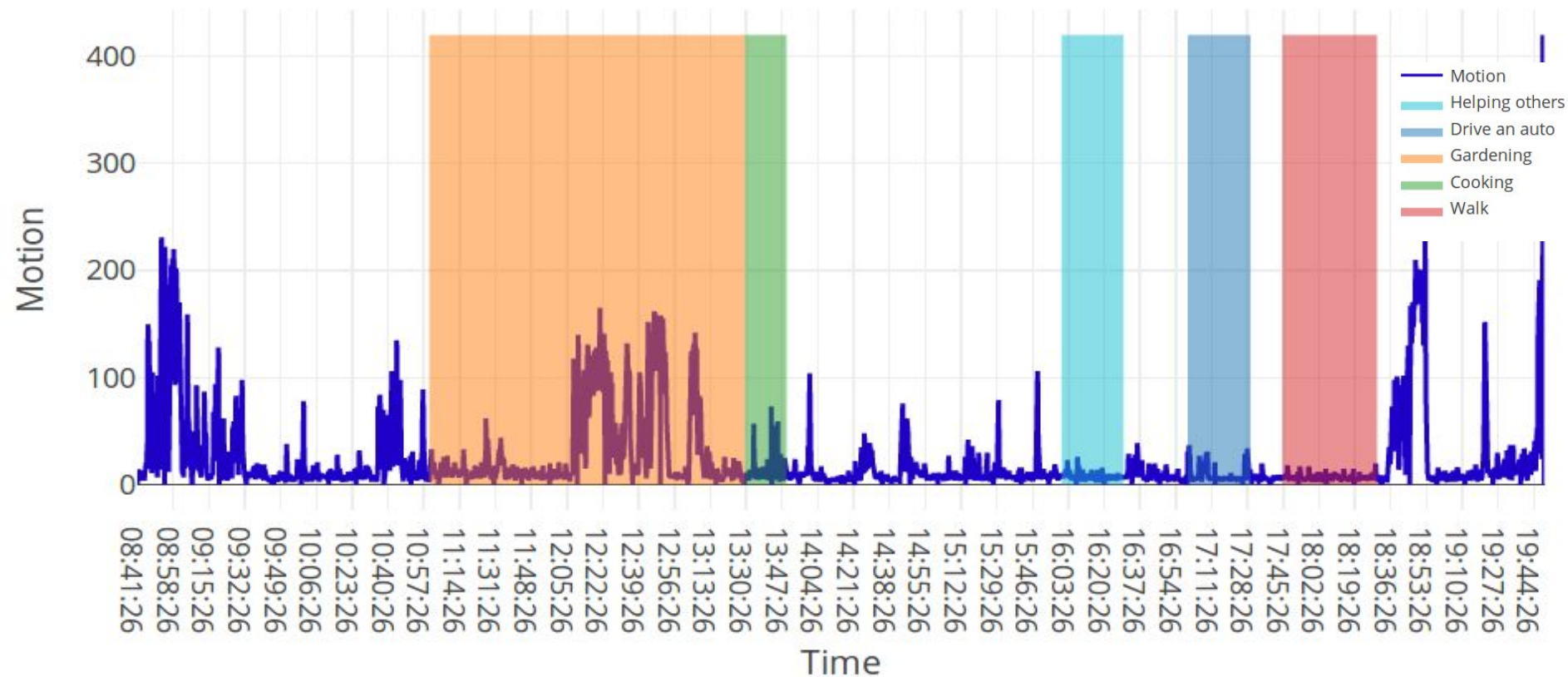
- Allows to determine the level of sweating
- Was used by ~300 volunteers
- During each day of studies volunteers were keeping track of activities



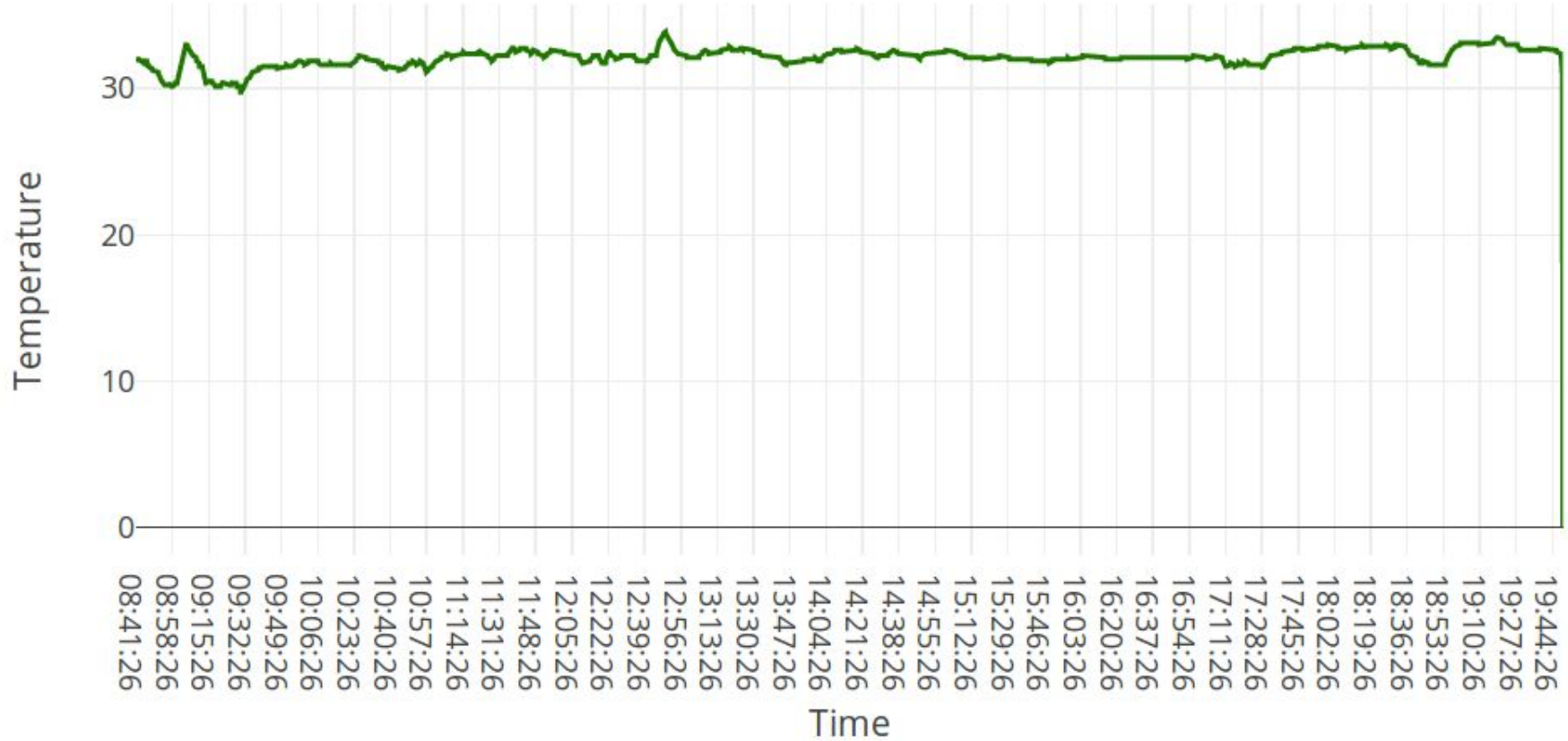
THE DATA: CONDUCTIVITY (SWEAT LEVEL)



THE DATA: MOTION



THE DATA: TEMPERATURE



ACTIVITIES

Stress sweaters:

Discussions or dispute
Had a date or heart palpitation
Had an embarrassing situation
Had too short / stressful lunch break
Ignored
Presentation or lecture or give lecture
Standing in a traffic jam
Standing in the queue
Stress when parking / parking space search
Stressful situations at work (difficult meeting, salary negotiation, visiting fair, etc.)
Unpleasant phone call
Visit doctor or dentist
Watched football match
Wear a dress

Activity sweaters:

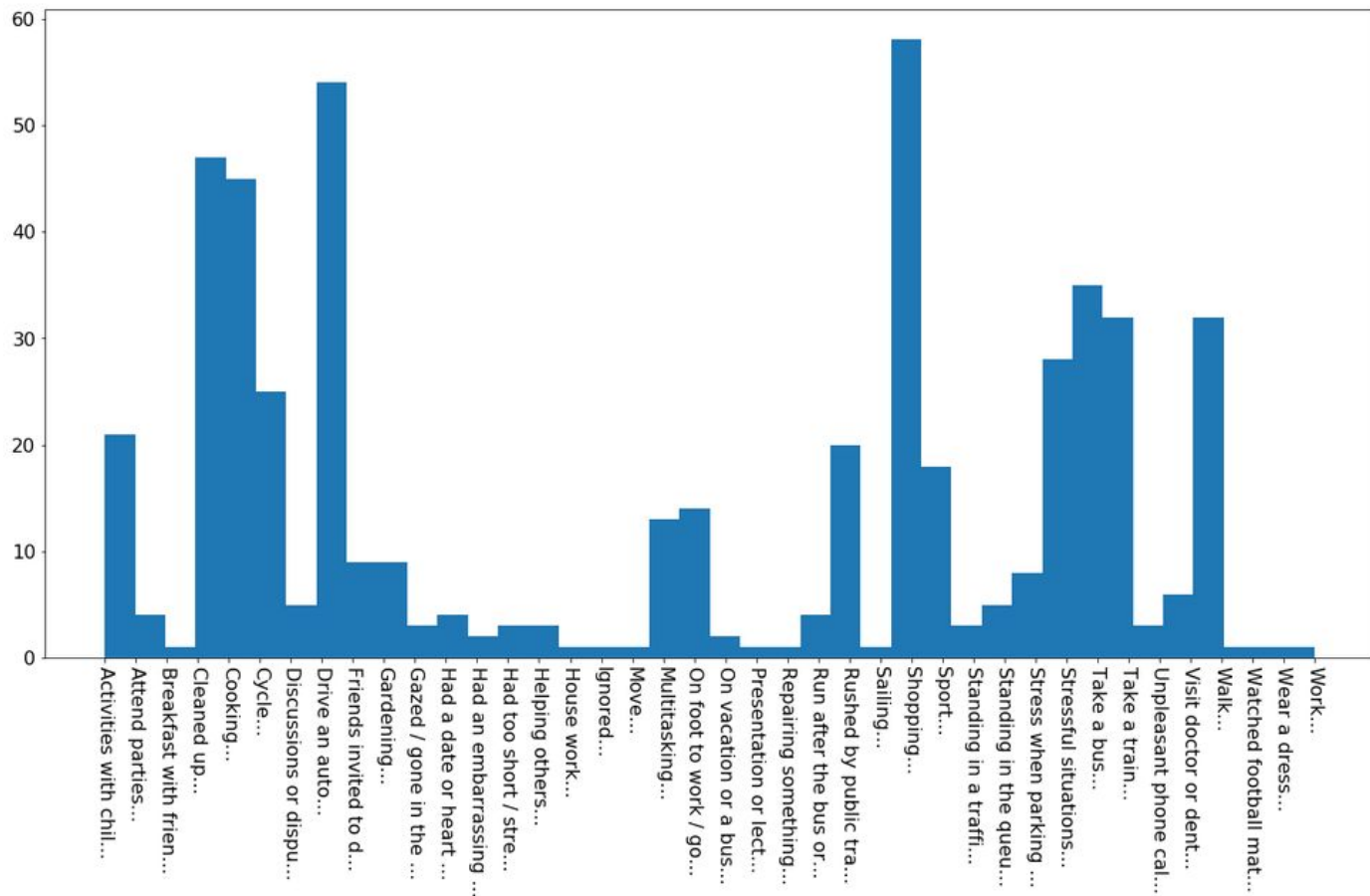
Cycle
Dancing, going out
Gardening
Move
On foot to work / gone home
Sport
Walk

Other:

Activities with children
Attend parties
Breakfast with friends
Cleaned up
Cooking

...

ACTIVITY TYPES



Sport vs Stress activities

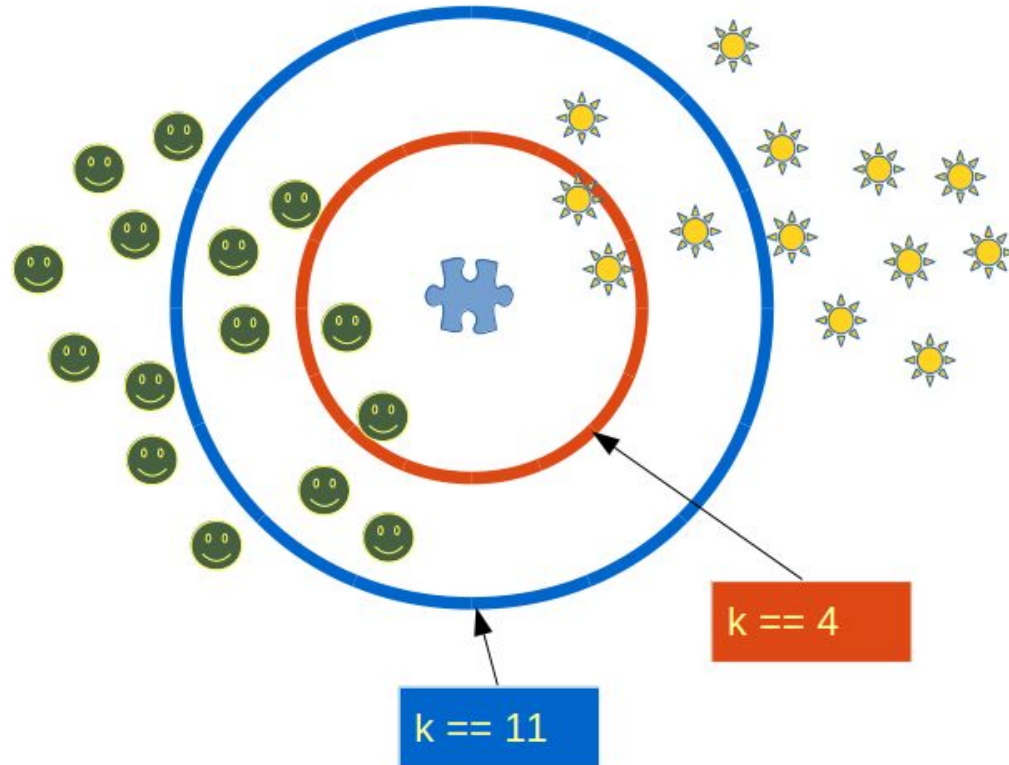
Your actions?

DATA CLEANING

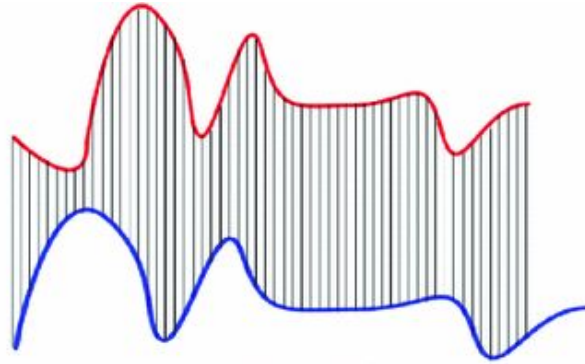
1. Make time intervals more precise?
2. Remove samples with constant values
3. Remove / smooth peaks

METRIC APPROACH: KNN

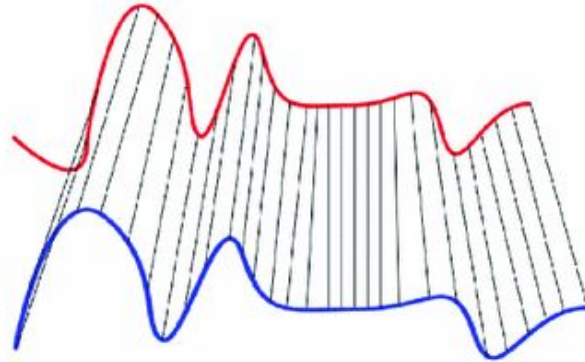
🧩 == 😊 or 🧩 == ☀️ ?



SELECTING DISTANCE FUNCTION

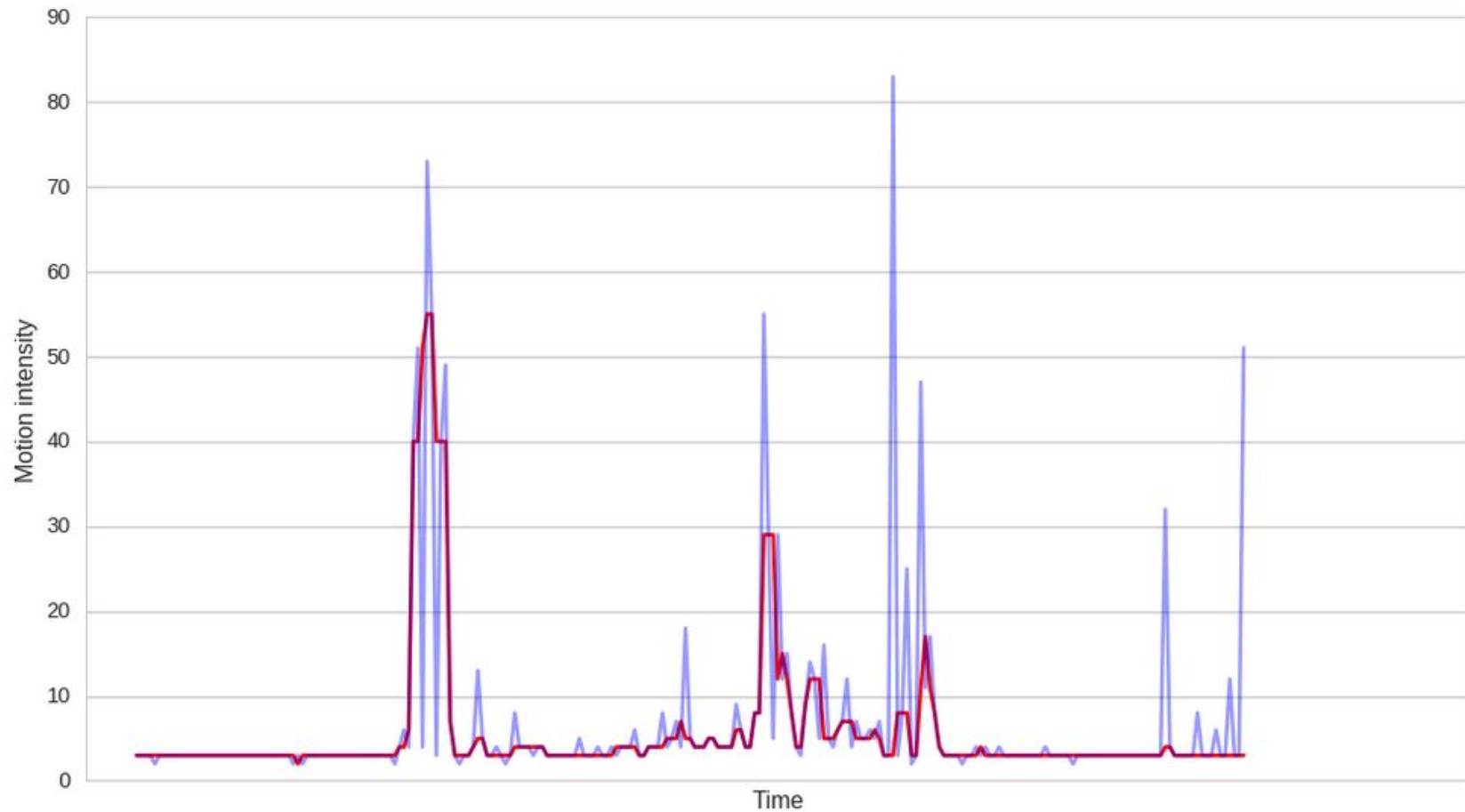


Euclidean Matching



Dynamic Time Warping Matching

SMOOTH RANDOM PEAKS



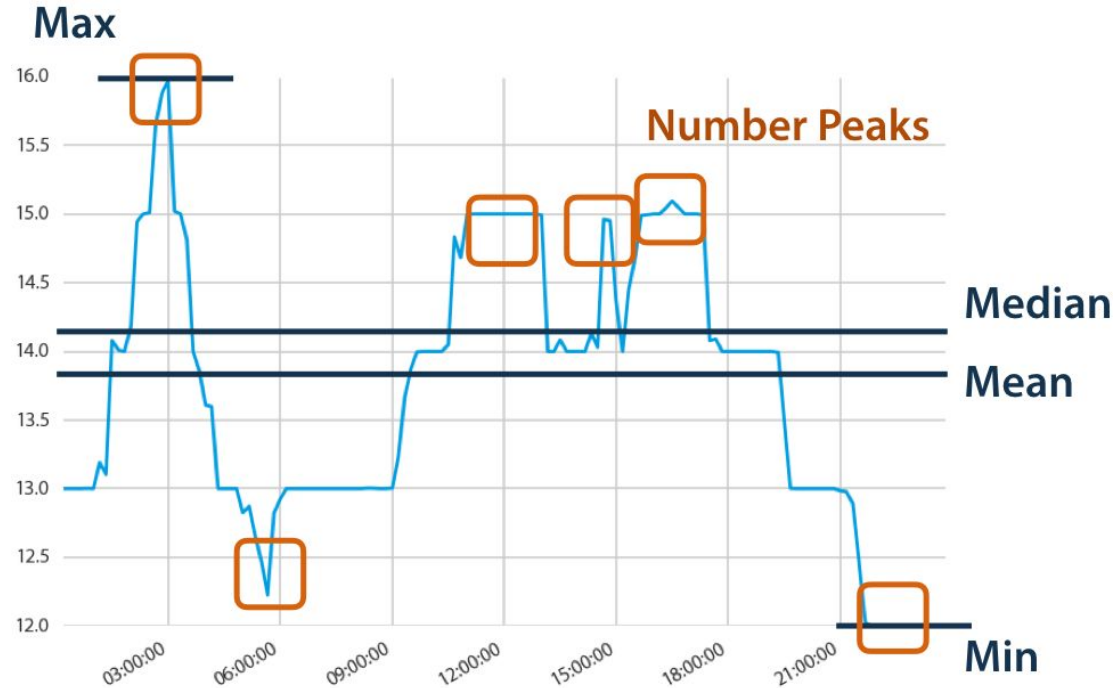
45% accuracy on small balanced dataset

Extremely slow

Imprecise labels

Your actions?

FEATURE EXTRACTION: TSFRESH



FEATURE EXTRACTION: TSFRESH

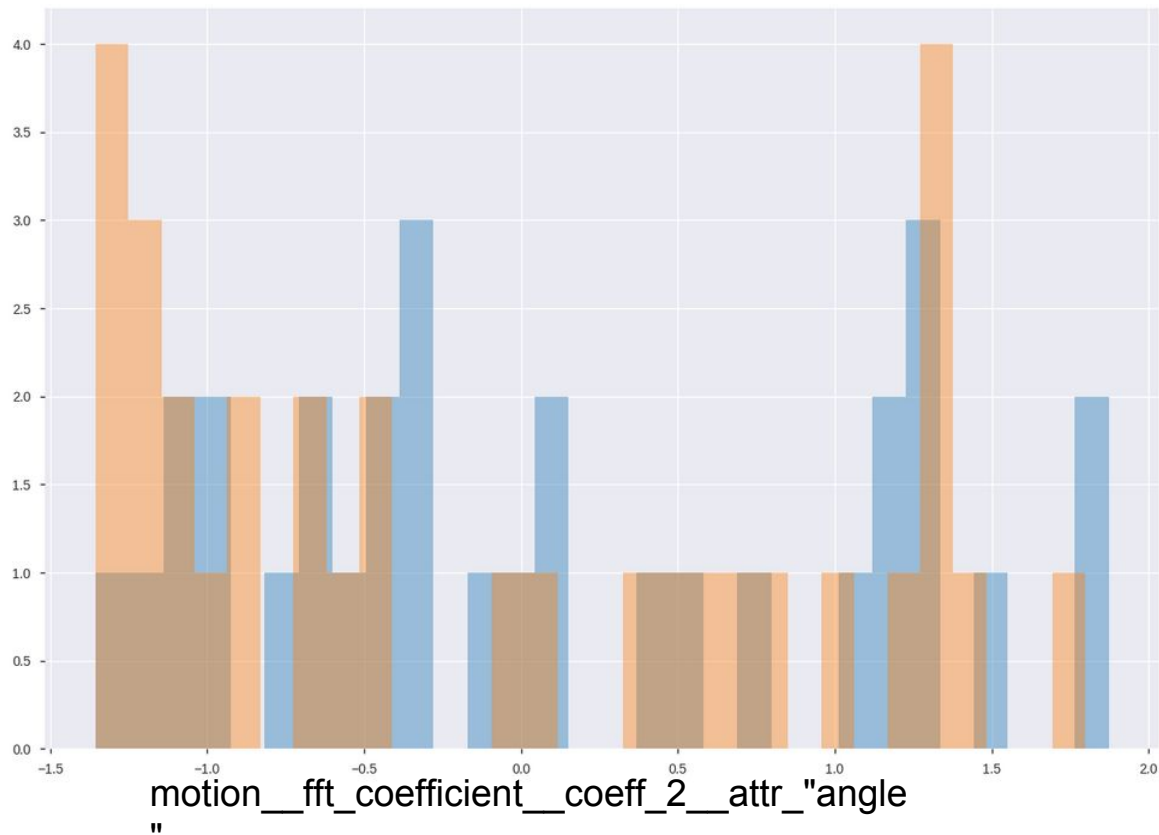
1 timeseries => 1000 features, 15000 features in total

1. Remove constant features
2. Apply recursive feature elimination
3. Remove linearly dependent features

EXPERIMENTS

Model	Accuracy, %
Logistic regression	96
Random forest	80
LightGBM	70

EXPERIMENTS



Bad features?
Classes are not separable?

1. Feets package: 57 features per time series
2. Features based on sweat event
3. Max, min, day of month, weekday, thresholds

1. Look at feature distribution, correlation
2. Look at feature importances from Tree-based classifiers
3. Look at feature coefficients from LogisticRegression

“ACTIVE” ACTIVITIES VS ALL

Active break

Climb stairs

Cycle

Move

On foot to work / gone home

Physical work

Play football

Sailing

Sport

Walk

ONE CLASS VS ALL

- **Cycle**
- **Cook**
- Drive an auto
- Stressful situations at work
- Activities with children
- Rushed by public transport during rush hour
- Cleaned up
- **Take a train**
- **Sport**
- Walk
- Multitasking
- Shopping
- Take bus
- **On foot to work / gone home**

ACTIVITY RECOGNITION: STATE OF THE ART

State-of-the-art research use:

1. Very limited set of classes (6 vs 30+)
2. High frequency of recording (50 Hz vs 0.1 Hz)
3. 3-axial sensors of motion and angles
4. Videos of people doing activities for precise manual labelling

and achieve 98% accuracy

In other words:

1. Better hardware
2. More accurate labels

REFERENCES

1. <https://github.com/markdregan/K-Nearest-Neighbors-with-Dynamic-Time-Warping>
2. <https://github.com/blue-yonder/tsfresh>
3. <https://github.com/carpyncho/feets>