CSE 118-218

Behavioral Context Recognition

Yonatan Vaizman

UCSD

Fall 2017

Outline

1. Introduction

   What is a context-recognition system?

2. Sensors & contexts

   What is the input and output of the system?

3. AI & ML

   What's the mechanism inside the system? And how to train it?

4. Data & validation

   Does the system work? Let's check with actual data!

5. Applications

   What can we do with the system?

   - Your class project

trade-offs, ExtraSensory

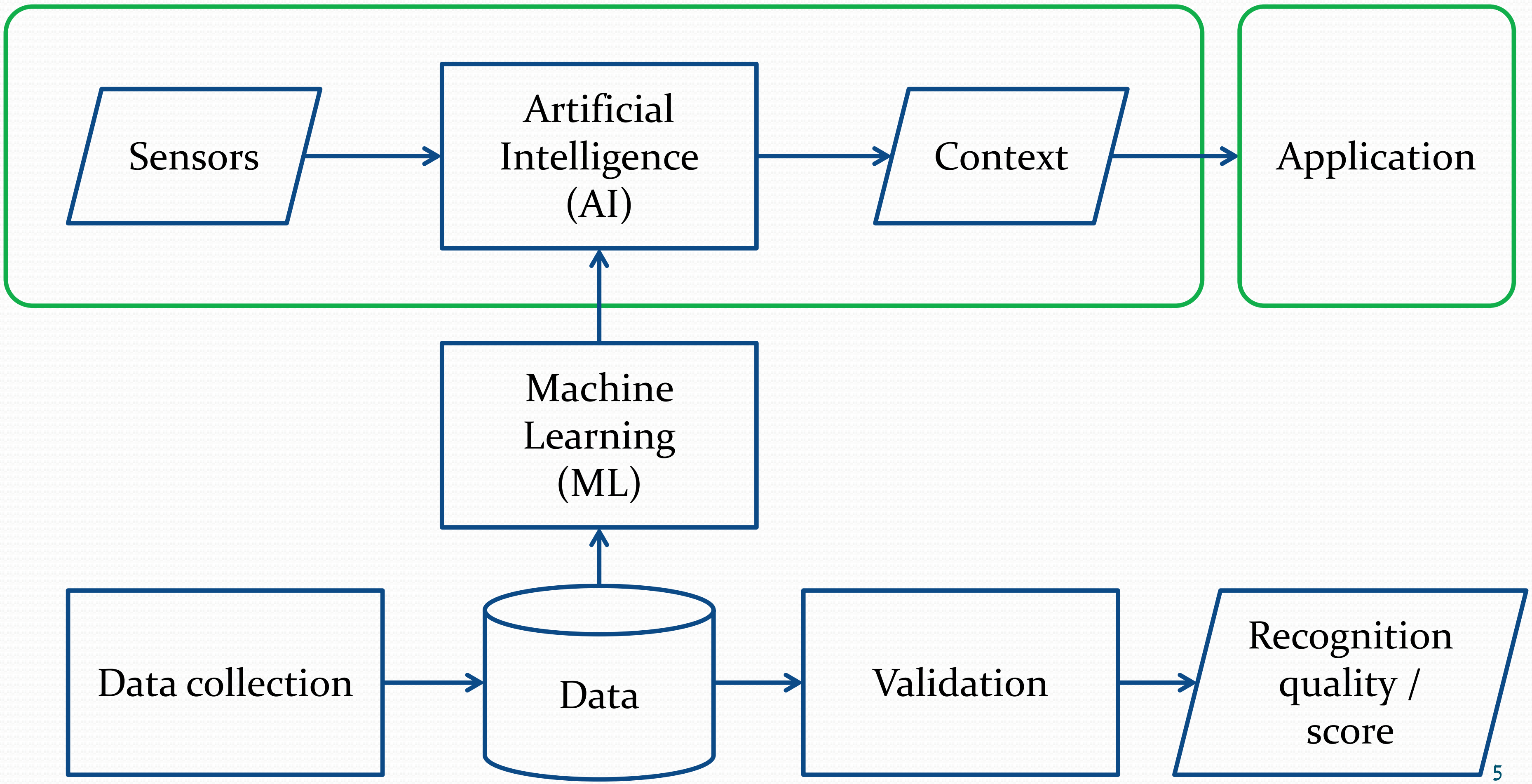
1. Introduction

What is a context-recognition system?

Motivation

- Music recommendation. Gert Lanckriet lab (ECE)
  - Context-awareness
- Public health. Kat Ellis
  - Physical activity (sitting, running, lifting weights…)
  - Social activity (business meeting, date, hang out…)
  - Functional independence (shower, cooking, cleaning…)
  - Environment (home, office, beach, restaurant…)
  - Emotions/mood
  - …
- Behavioral context recognition

Context Recognition System

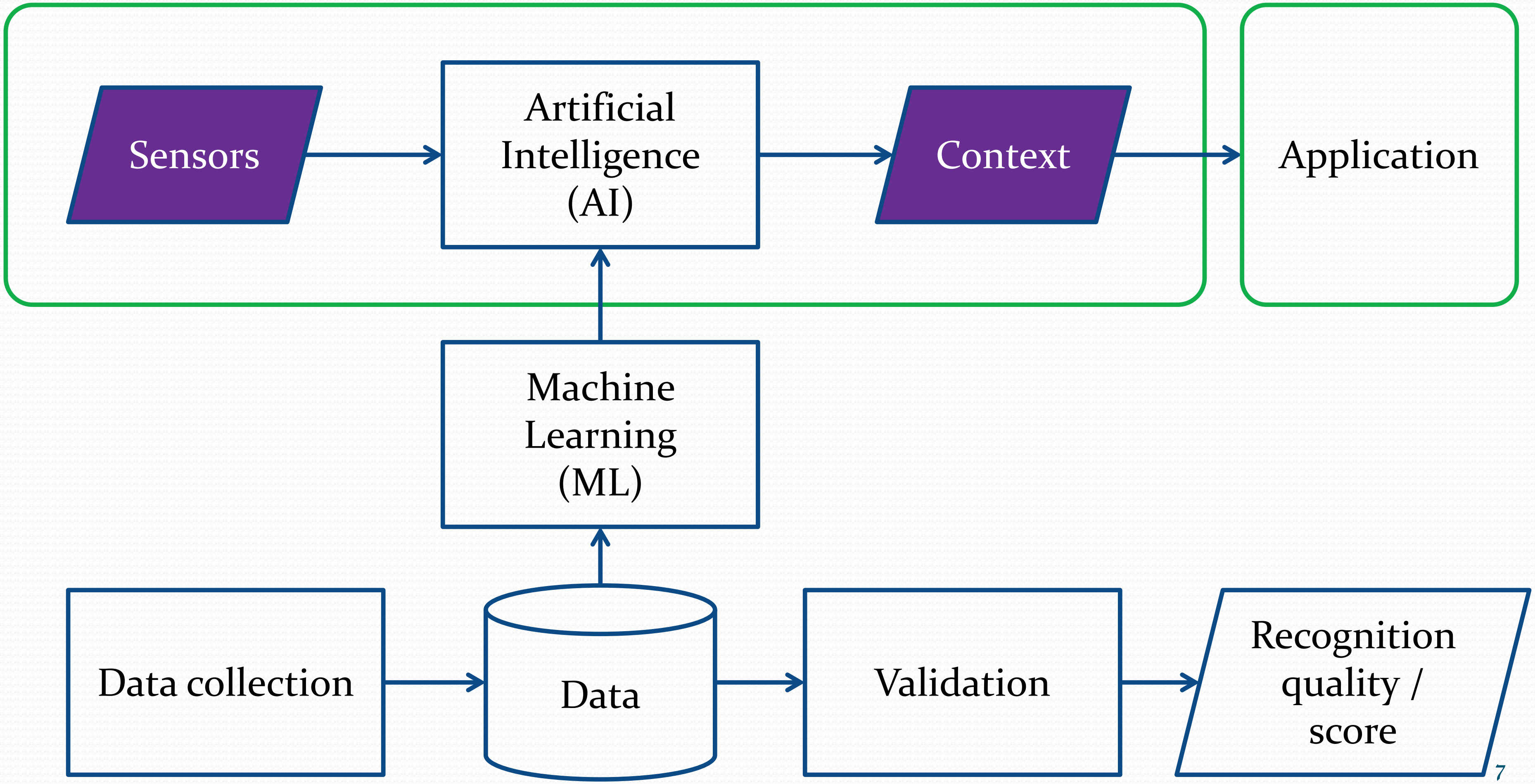


2. Sensors & contexts

What is the input and output of the system?

2. Sensors & contexts

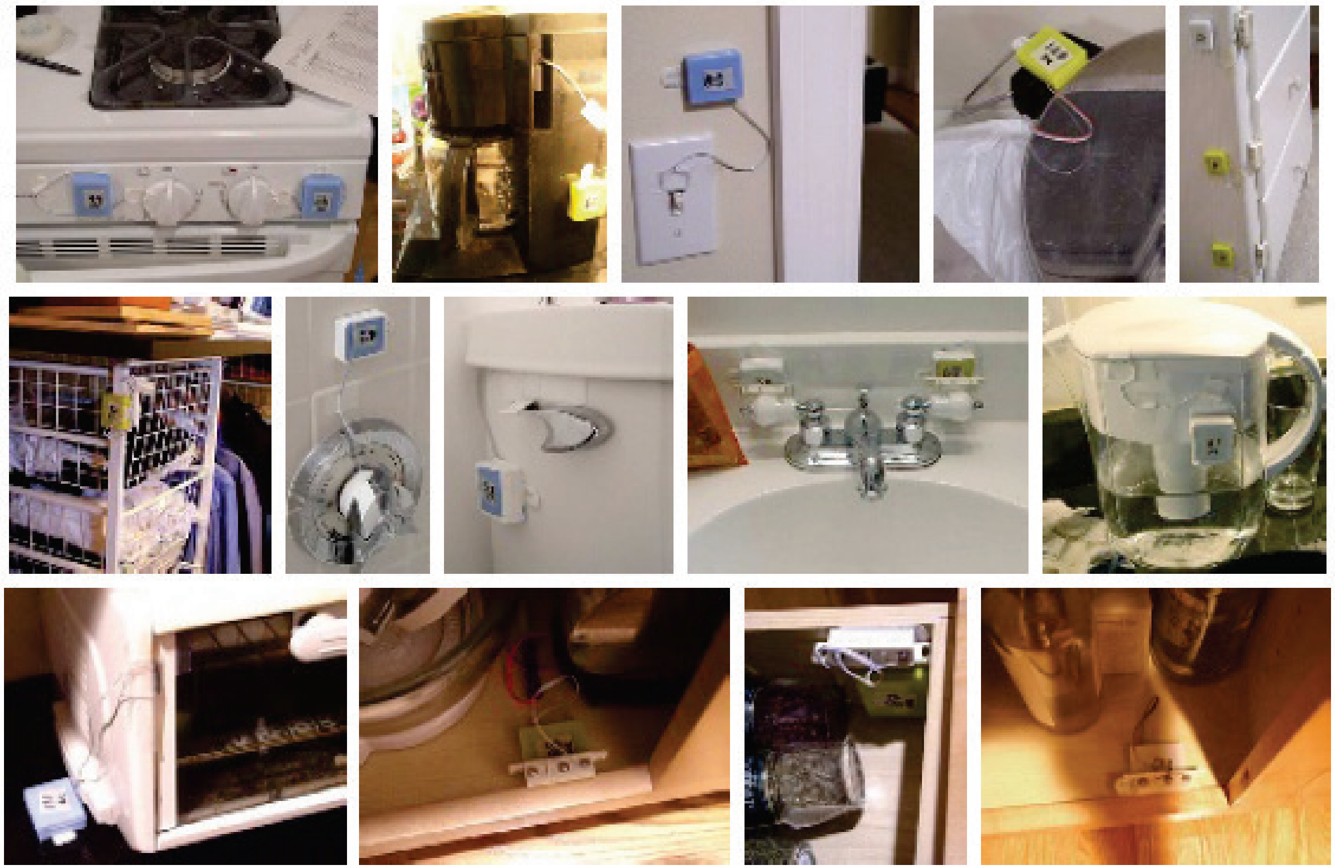
What is the input and output of the system?



Sensing modality:

Stationary – state-change

Tapia, 2004



Recognized contexts:
Preparing breakfast
Preparing lunch
Bathing
Doing laundry
Going to work
Watching TV
…

Possible applications?

Sensing modality:

Acceleration

Bao, 2004

Recognized contexts:
Walking
Sitting & relaxing
Standing still
Watching TV
Running
Folding laundry
Brushing teeth
Riding elevator
Vacuuming
Climbing stairs
…

Possible applications?

Sensing modality:

Audio

Yatani, 2012 (BodyScope)

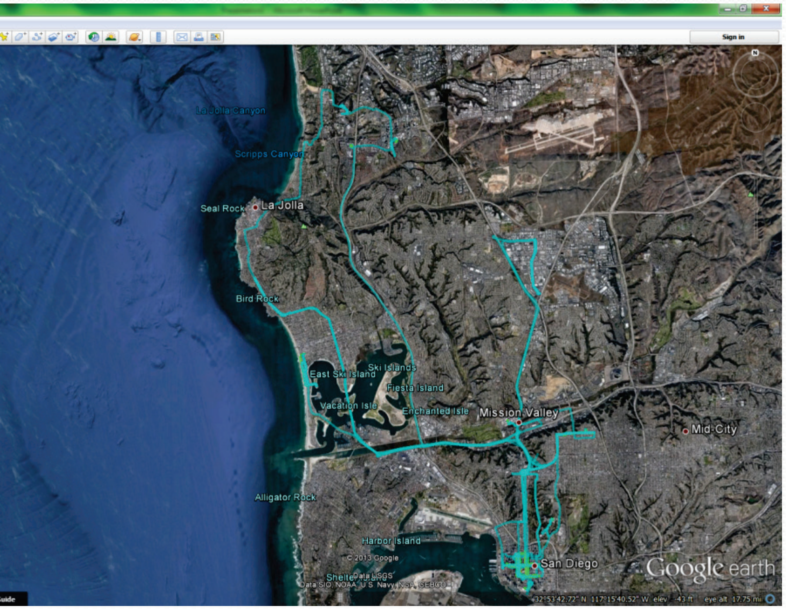
<u>Recognized contexts:</u>
Deep breath
Eating (cookie)
Eating (bread)
Drinking
Speaking
Whispering
Whistling
Caughing
…

Possible applications?

Sensing modality:

Location (GPS)

Ellis, 2014 (GPS + accelerometer)



<u>Recognized contexts:</u>
Sitting
Walking/Running
Bicycling
Vehicle
(Car vs. Bus)

Possible applications?

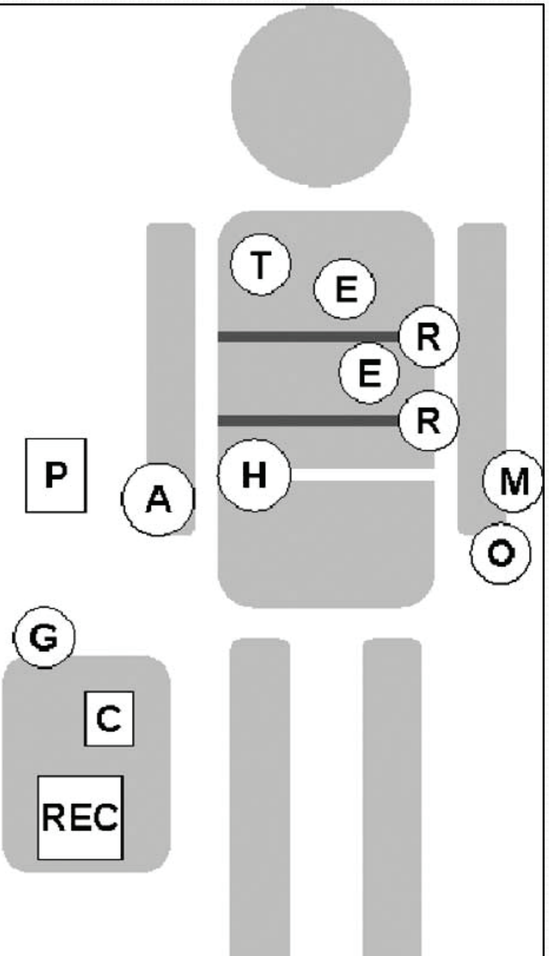
Multi-modal sensing

Ermes, 2008

Wearing:
A = 3D accelerometers on wrist
H = Sensorbox on hip containing 3D accelerometers, 3D magnetometers, environmental temperature, illumination, and humidity
T = Skin Temperature sensor
E = ECG electrode
R = Respiratory effort sensor
M = MP3-audio player/recorder
O = Oximeter

Rucksack:
G = GPS receiver
C = Camera
REC = 19 channel recorder

Manual annotation:
P = PDA



Recognized contexts:
Lie
Sit
Stand
Walk
Run
Bike
Exercise-bike
Rowing machine
…
Indoors vs. outdoors
Eating vs. not eating
…

What do you think about this sensor setup?

Practical devices

- Mobile Sensing Platform (MSP)
  - Unobtrusive: single small device
  - Multi-modal sensors
  - Wireless

Choudhury, 2008

Sensor description	Maximum sampling rate used
Electret microphone	16,384 kHz
Visible light phototransistor	550 Hz
3-axis digital accelerometer	550 Hz
Digital barometer temperature	30 Hz
Digital IR and visible+IR light	5 Hz
Digital humidity/temperature	2 Hz
Digital Compass	30 Hz

How about everyday devices? Unobtrusive, multi-modal, and natural to use?

Smartphone & Smartwatch

Guiry, 2014

<u>Phone-sensors:</u>
Accelerometer
Magnetometer
Gyroscope
Air pressure

<u>Watch-sensors:</u>
Accelerometer

<u>Recognized contexts:</u>
Walking
Running
Cycling
Standing
Sitting
Elevator ascent
Elevator descent
Stair ascent
Stair ascent

<u>Phone-sensors:</u>
GPS
Light

Indoors vs. outdoors

ExtraSensory

Vaizman, 2017

<u>Phone-sensors:</u>
Accelerometer
Magnetometer
Gyroscope
Audio
Location (GPS, WiFi, cellular)
Phone-state

Light
Air pressure
Humidity
Temperature
Proximity

<u>Watch-sensors:</u>
Accelerometer
Compass

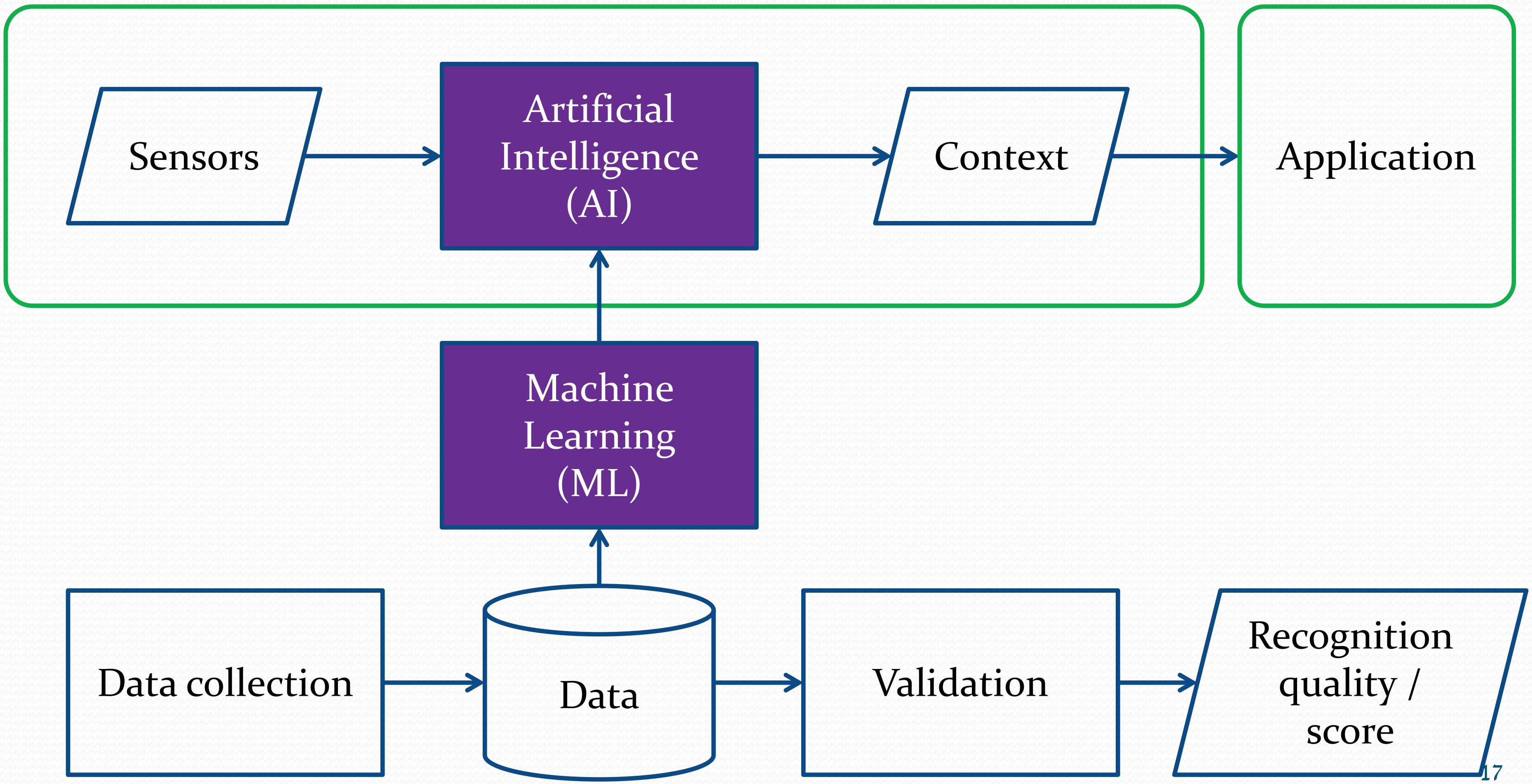
<u>Recognized contexts</u>
<u>(**Multi-label**):</u>
Walking
Running
Bicycling
…
Indoors
At school
At home
At the beach
At a bar
…
Sleeping
Eating
Toilet
Shower
Watching TV
…
Driving
On a bus
…

3. AI & ML

What's the mechanism inside the system? How to train it?

3. AI & ML

What's the mechanism inside the system? How to train it?



Artificial Intelligence

- Machine performing intelligent task

- Imitate human cognitive function:
  - Speech recognition
- Beyond human:
  - Play chess (really well)
- Non human:
  - GPS – infer global location from satellite radio signals

AI approaches

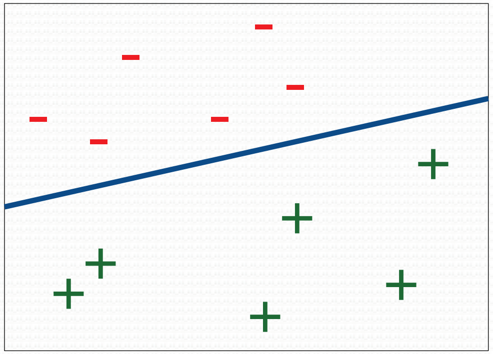
- Explicit: rule-based / formula-based
  - Geolocation (GPS)
- Implicit: data-based

- Learning from examples

Machine Learning (supervised)

- Classifier: $h: \mathcal{X} \rightarrow \mathcal{Y}$
- Feature (input) space $\mathcal{X}$ (e.g. $\mathbb{R}^2$)
- Label (output) space $\mathcal{Y}$ (e.g. $\{-, +\}$)
- Hypothesis space $\mathcal{H} \subseteq \{h: \mathcal{X} \rightarrow \mathcal{Y}\}$

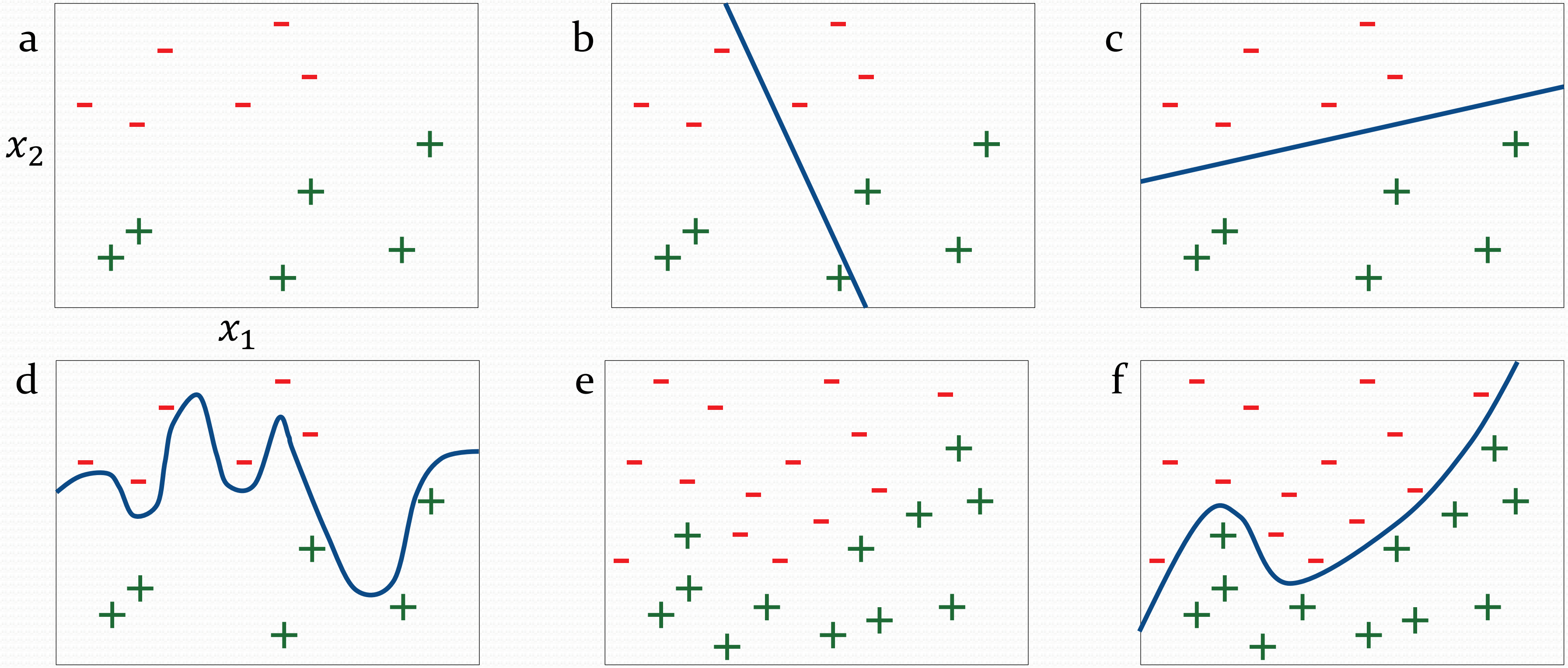
(e.g. linear separators)



- Training sample (data) $S = \{x_i \in \mathcal{X}, y_i \in \mathcal{Y}\}_{i=1}^{N}$

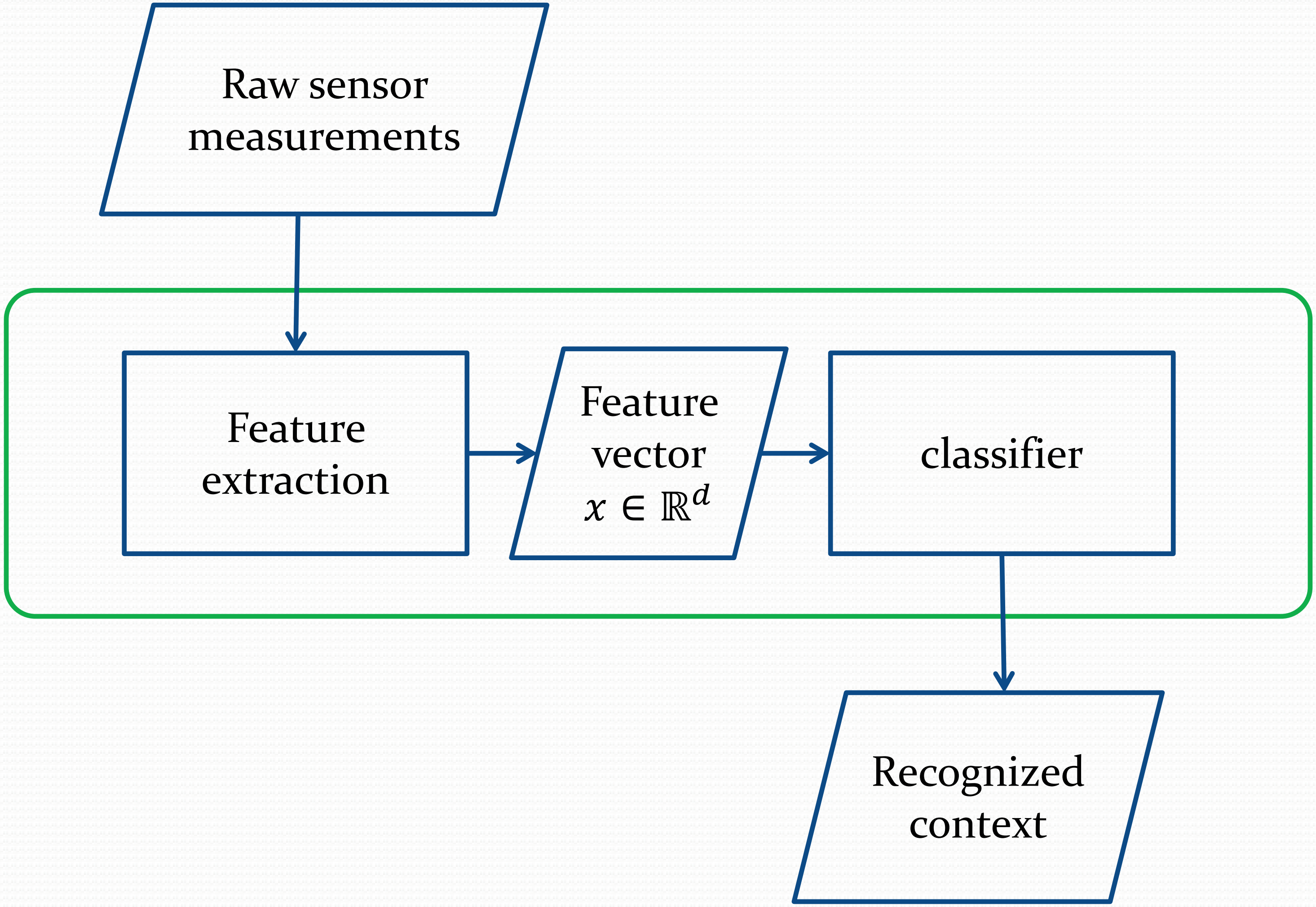
```
function learning_algorithm(sample S) {
        ...
        return classifier h*
}
```

Machine learning



- Fit the data well (c better than b)
- Avoid over-fitting to data (c better than d)
- With more data you can afford a more complicated rich model

Context recognition: Typical AI flow



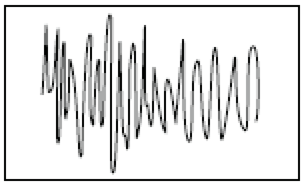
Feature extraction

*feature-learning (e.g. deep learning with plenty of data)

Vaizman, 2017 (October)

features

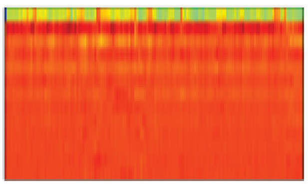
accelerometer



x11
x12
.
.
.

Magnitude statistics: mean, std, quantiles…
Spectral features (Fourier)
3d axes cross-correlations

audio



x21
x22
.
.
.

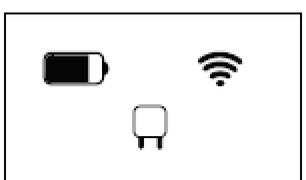
Mel Frequency Cepstral Coefficients (MFCC)
Mean, std MFCC

location

x31
x32
.
.
.

Variability of location:
std(latitude), std(longitude), distance travelled, average speed…

phone state



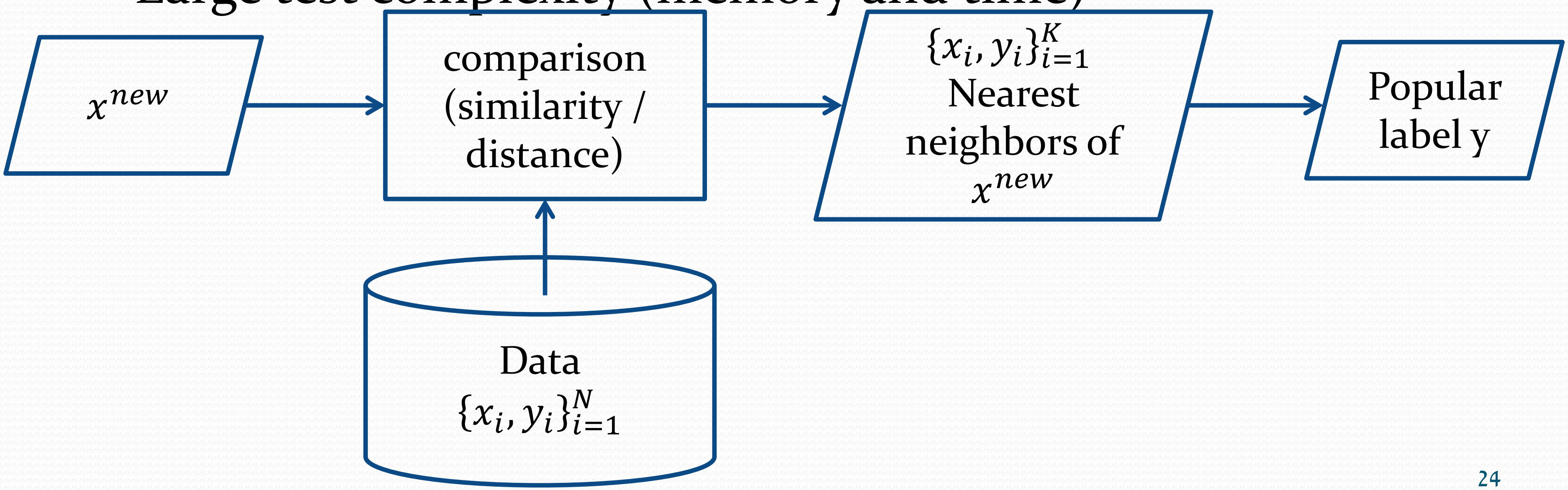
x41
x42
.
.
.

Binary state-indicators: WiFi available, app-foreground, battery charge state…

Instance-based classifier

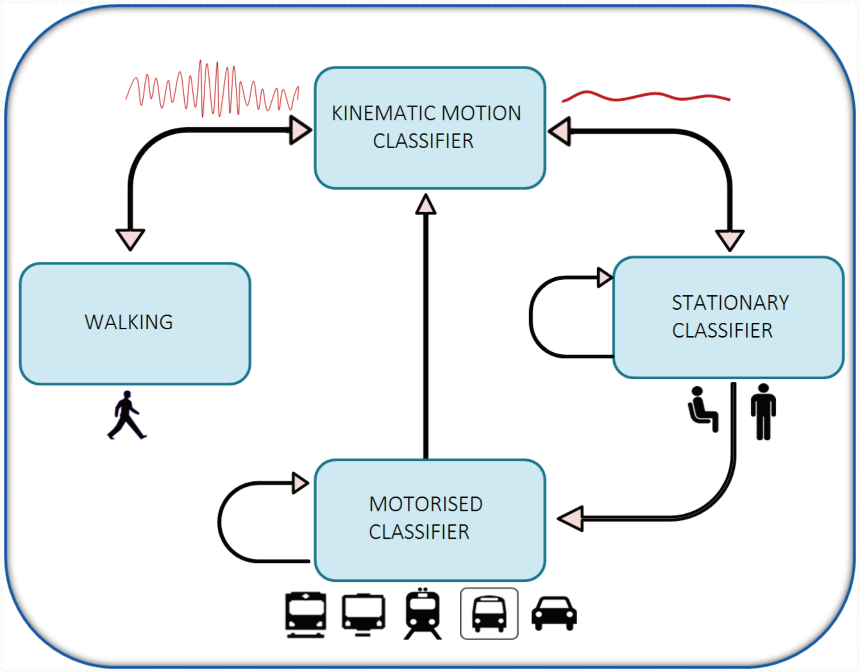
k-nearest-neighbors

- Strength:
  - Simple (to implement, understand, no training required)
- Weakness:
  - Assumes to find similar example
  - Large test complexity (memory and time)

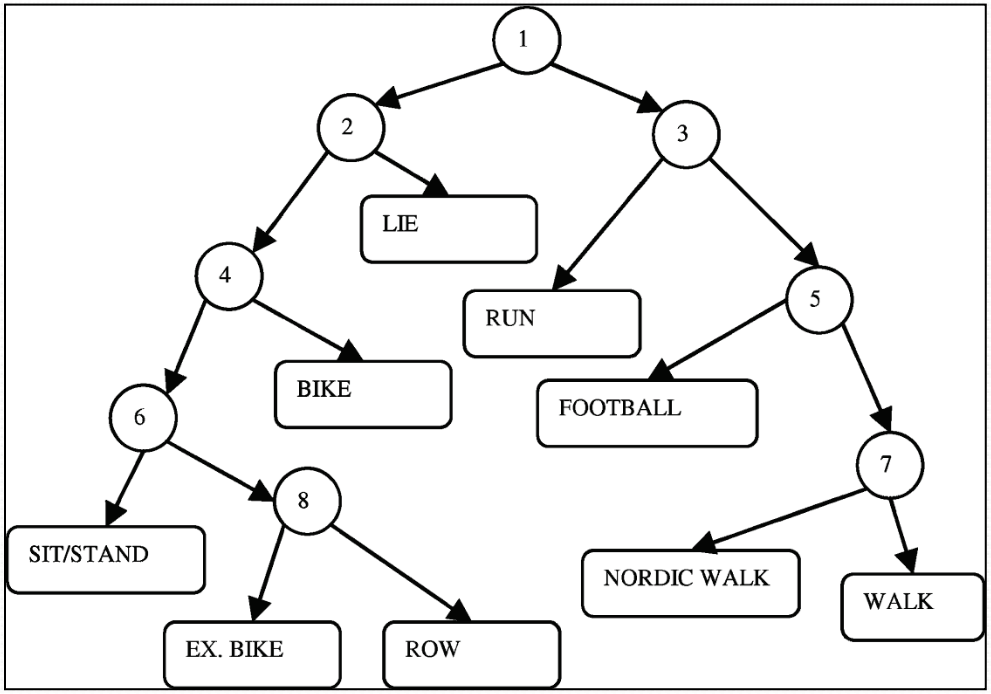


Task-designed classifier

Hemminki, 2013



Ermes, 2008



- Strength: Can server well a particular application
- Weakness:
  - Perhaps relies too much on researcher assumptions
  - Hard to generalize or scale to more behavioral contexts

Generic classifier

- Logistic regression, SVM
- Random forest
- Neural networks, Multiple-layer-perceptron (MLP)
- Time-series models, HMM
- …

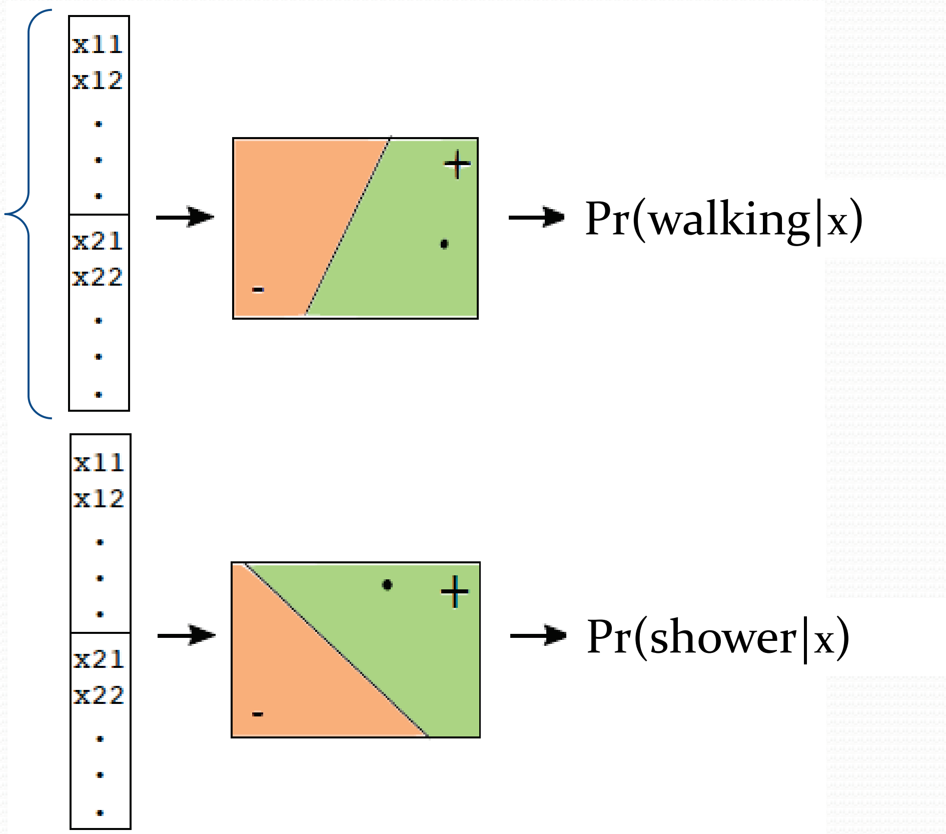
ExtraSensory

Logistic-regression per-label

Separate model per-label

Vaizman, 2017 (October)

175 features from 6 sensors

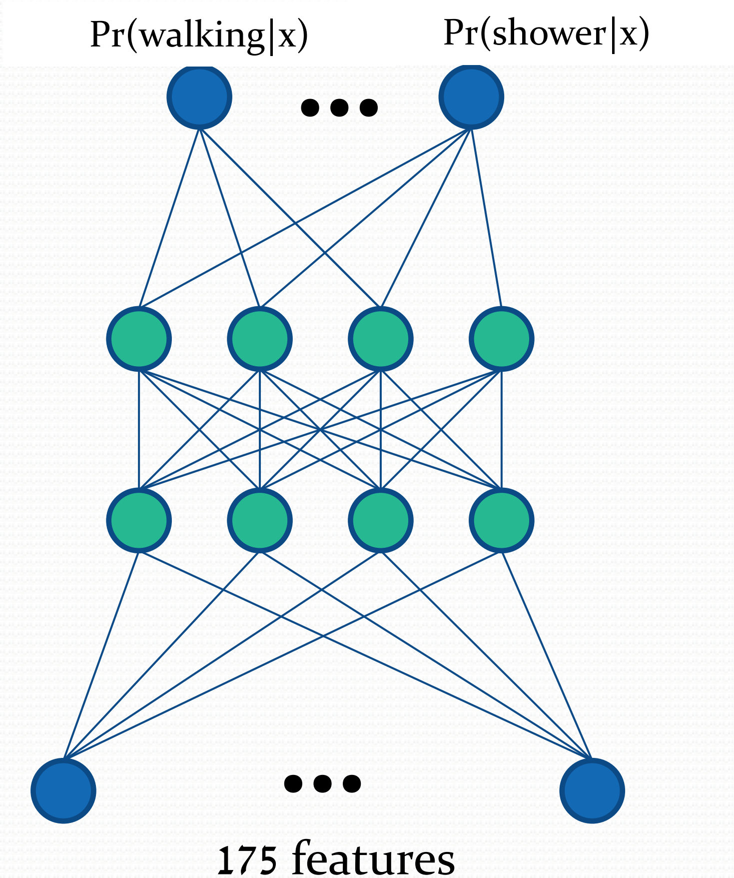


ExtraSensory

Multi-layer perceptron (MLP)

Shared model for 51 context labels

Vaizman, 2017 (December)



Predicted multi-label context

↑

Hidden representation

↑

Hidden representation

↑

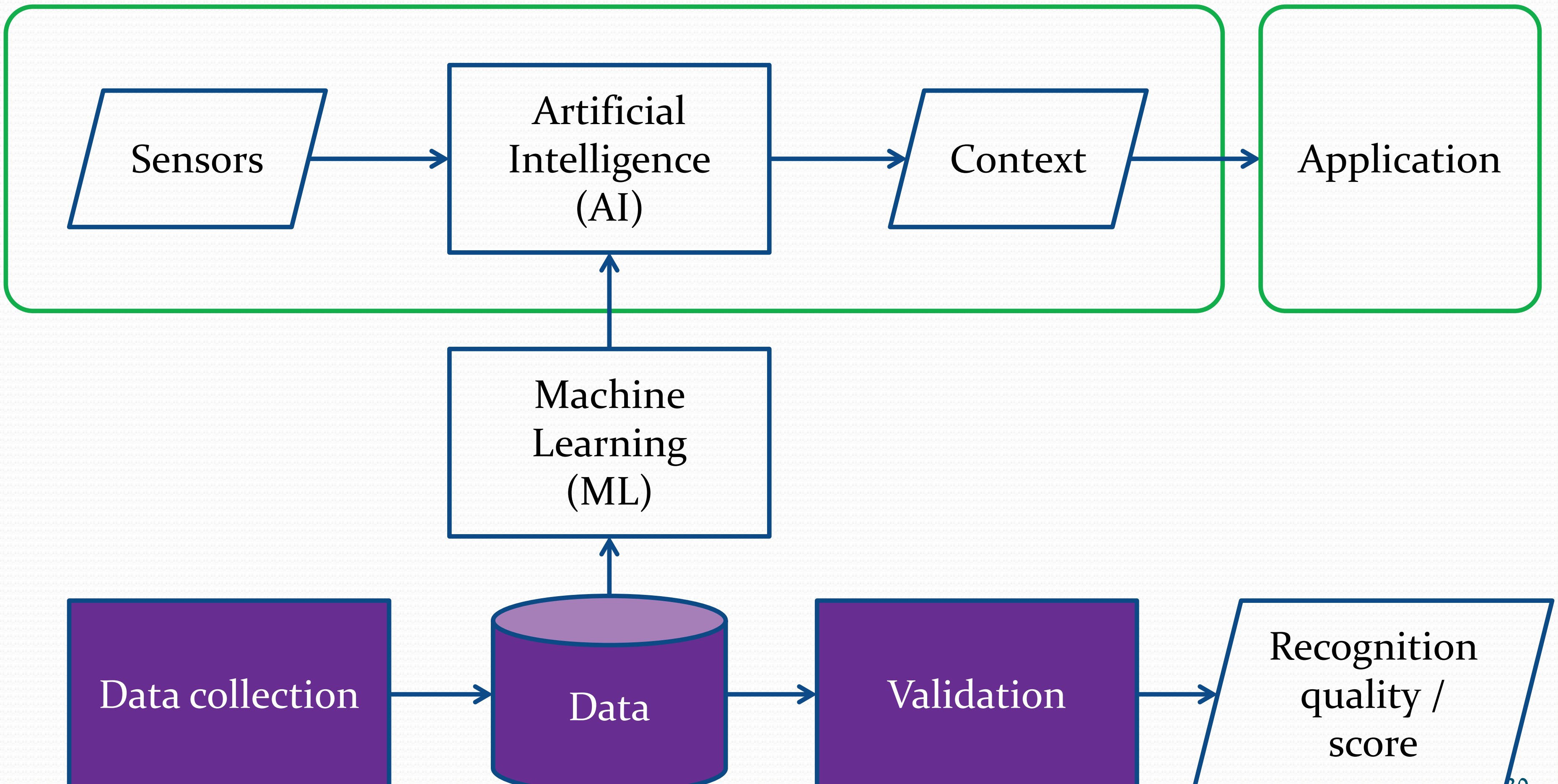
Input: x

4. Data & validation

Does the system work? Let's check with actual data!

4. Data & validation

Does the system work? Let's check with actual data!



Lab data

- Scheduled time, designated location
- Constrained device positioning
- Instructed, repeated, observed activities

- Strength:
  - Clean data: consistent, well-balanced
  - Reliable labels
- Weakness:
  - Doesn’t reflect richness of real-life behavior
  - Classification may be too easy (may generalize poorly)

Out of the lab

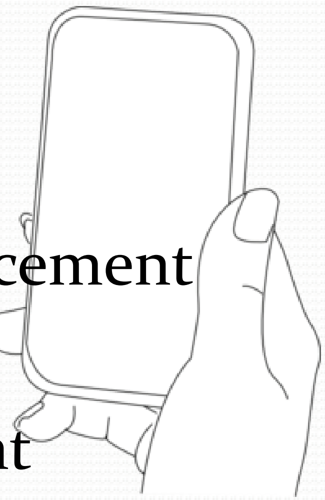
Ganti, 2010
Khan, 2014

- More natural behavior:
  - Natural environment, participant's own schedule
  - Unobserved, uninstructed behavior
- But still:
  - Prescribed targeted activities
  - Research device
  - Controlled device placement

In-the-wild

Vaizman, 2017 (October)

- ExtraSensory (2017):
  - Naturally used devices
  - Unconstrained device placement
  - Natural environment
  - Natural behavioral content



Multi-label

Walking, **Running**, Sitting, Standing, Lying down, Working, At home, At work, **At the beach**, Gym, **Listening to music**, Eating, Skateboarding, limping, Cleaning, Cooking, **Outside**, Indoors…

What are the difficulties/challenges in data collection in-the-wild?

Getting **labels** in-the-wild

- Without instructor
- Without observer
- How to get ground truth labels of behavioral context?

Getting labels in-the-wild:

researcher annotators

- Camera
  - Reliable labels, but harms privacy and uncomfortable equipment

Ellis, 2014

Pirsiavash, 2012

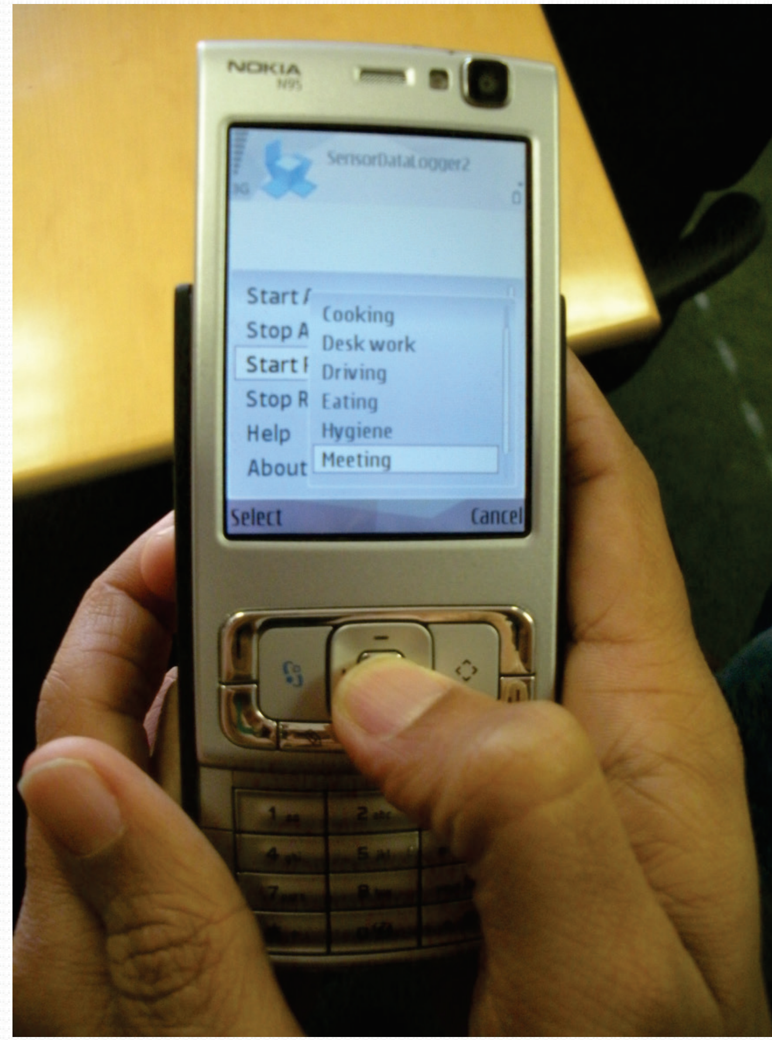
Getting labels in-the-wild:

self-reporting in-situ

- Simple labeling interface

Ganti, 2010

Ermes, 2008

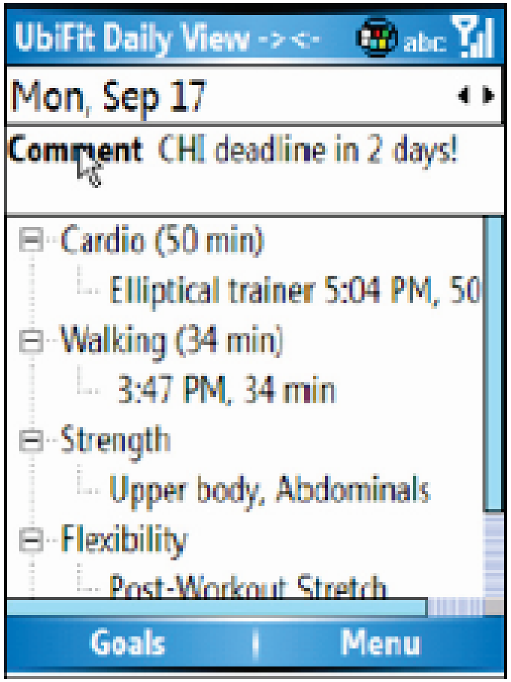


Getting labels in-the-wild:

self-reporting by-recall

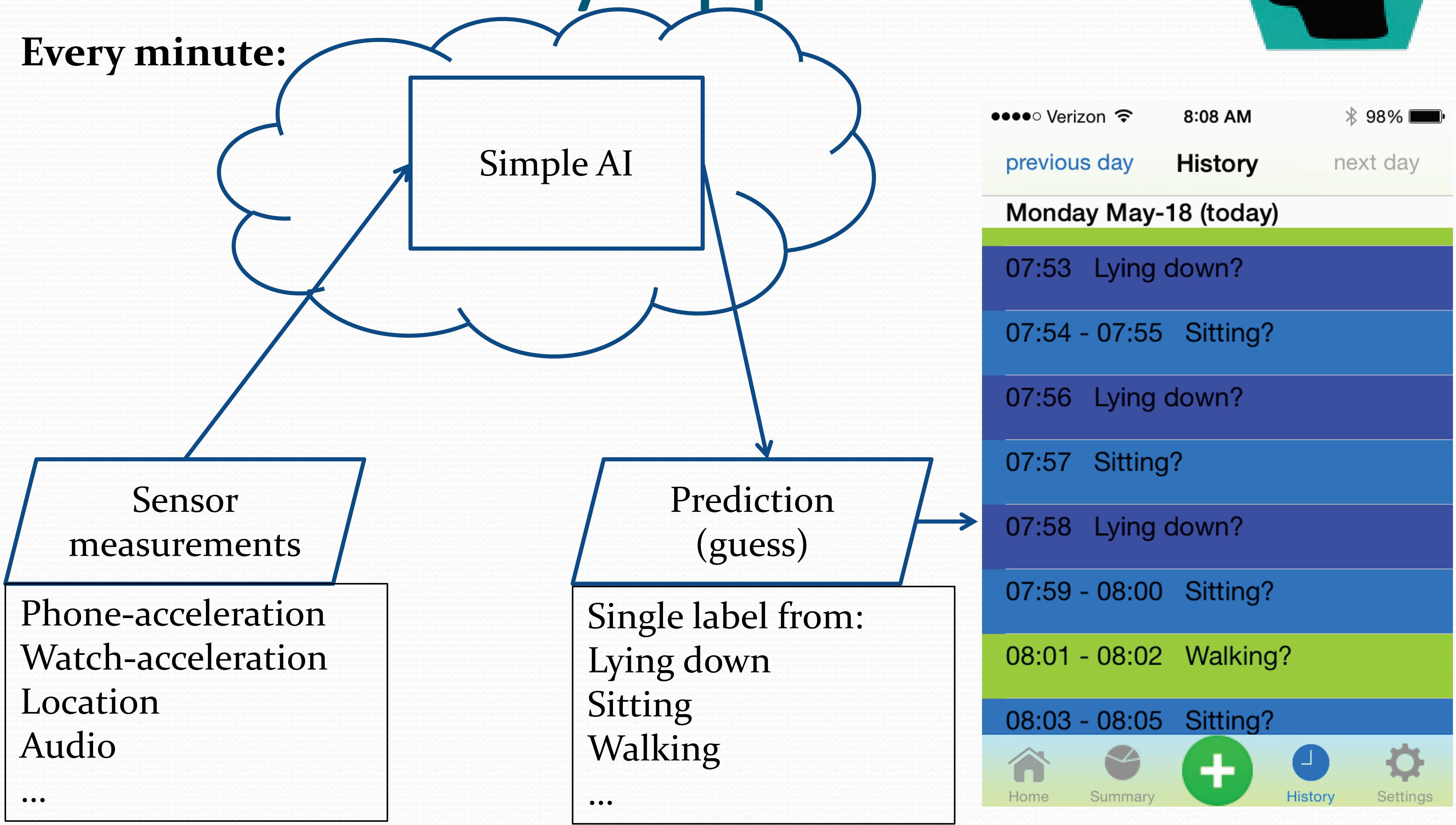
- Automatic recognition + manual labeling

Consolvo, 2008



ExtraSensory App

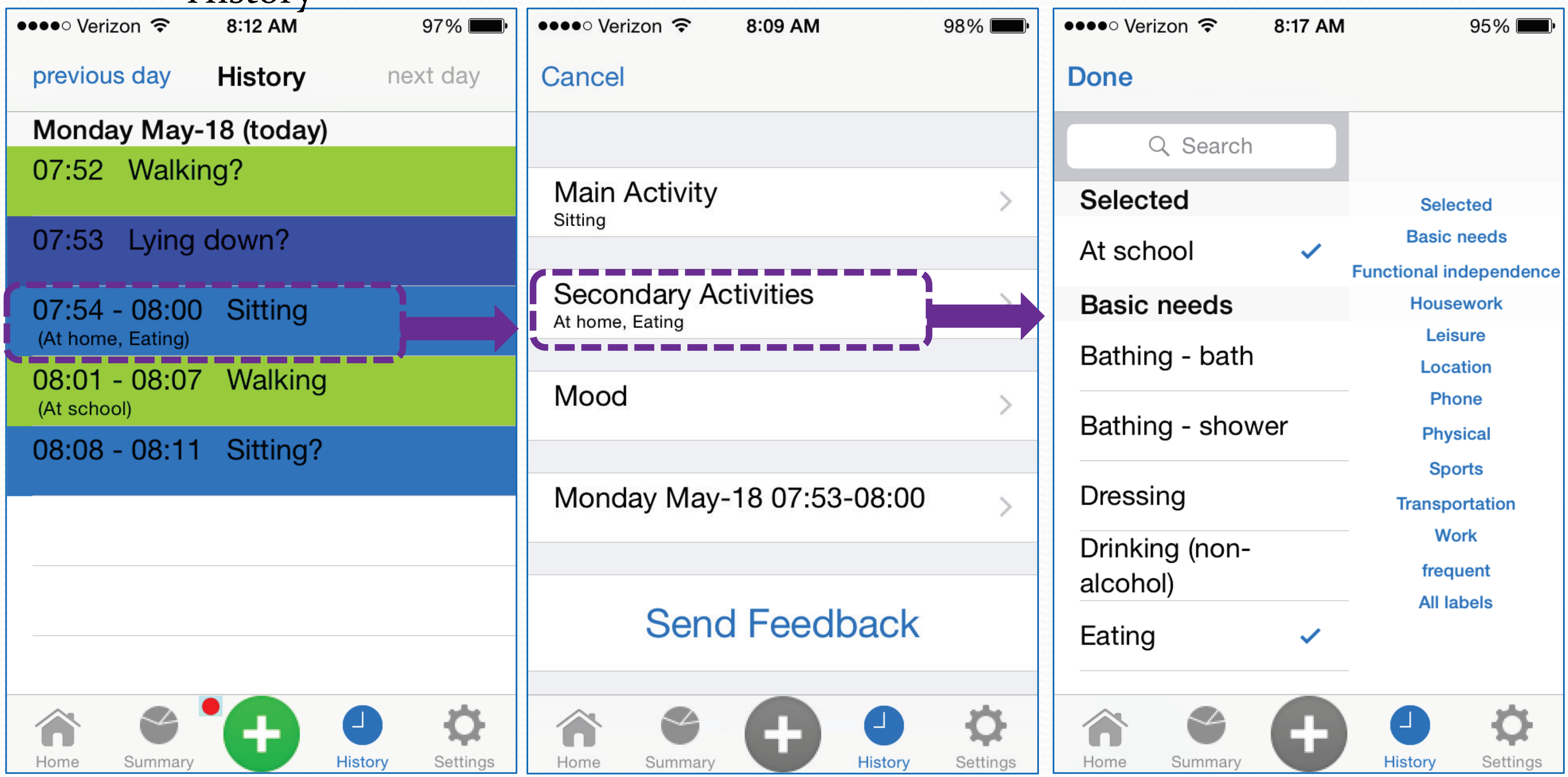
Every minute:



Getting labels in-the-wild:

Self-reporting with ExtraSensory App

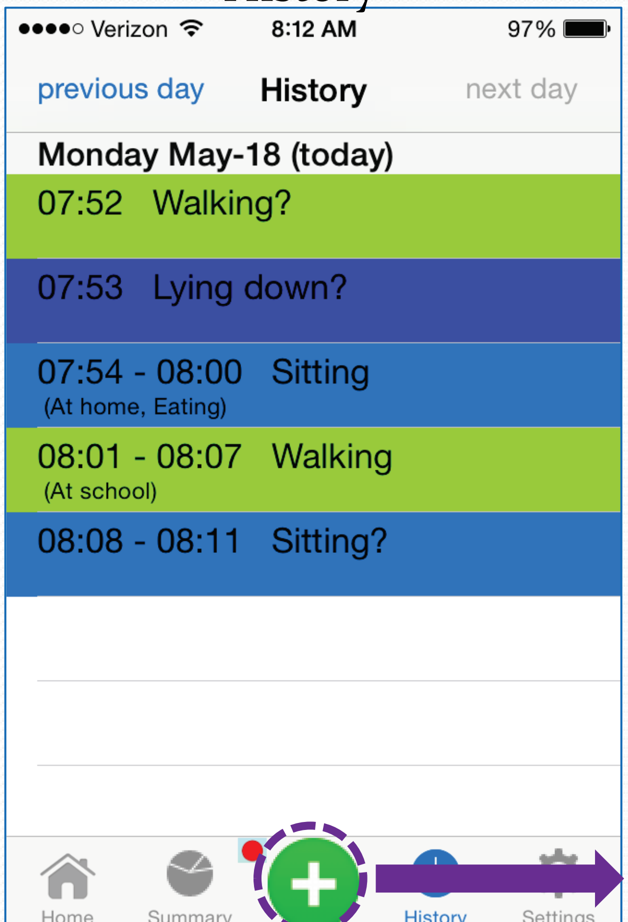
History



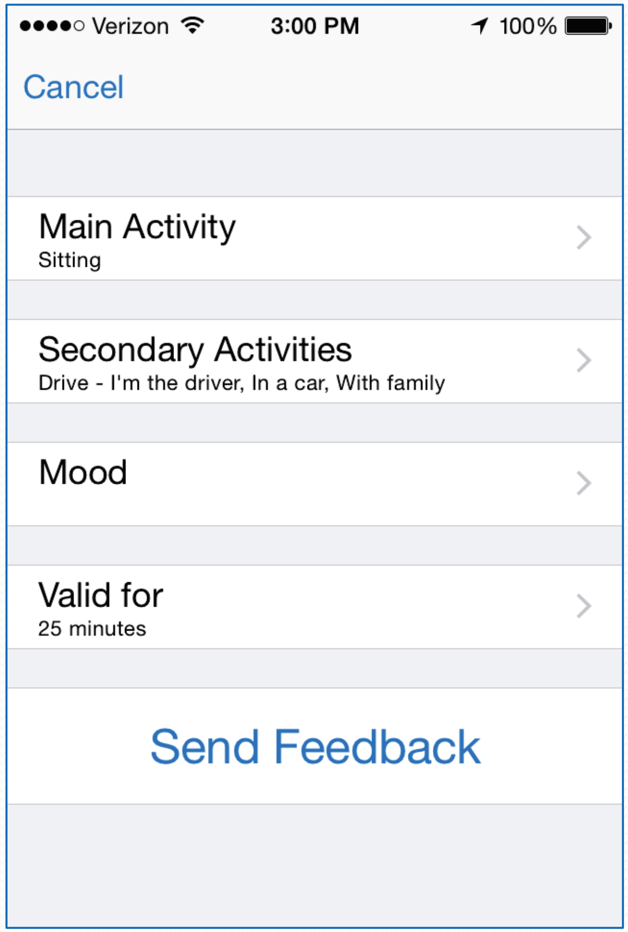
Getting labels in-the-wild:

Self-reporting with ExtraSensory App

History



Active feedback



Notification

ExtraSensory Dataset

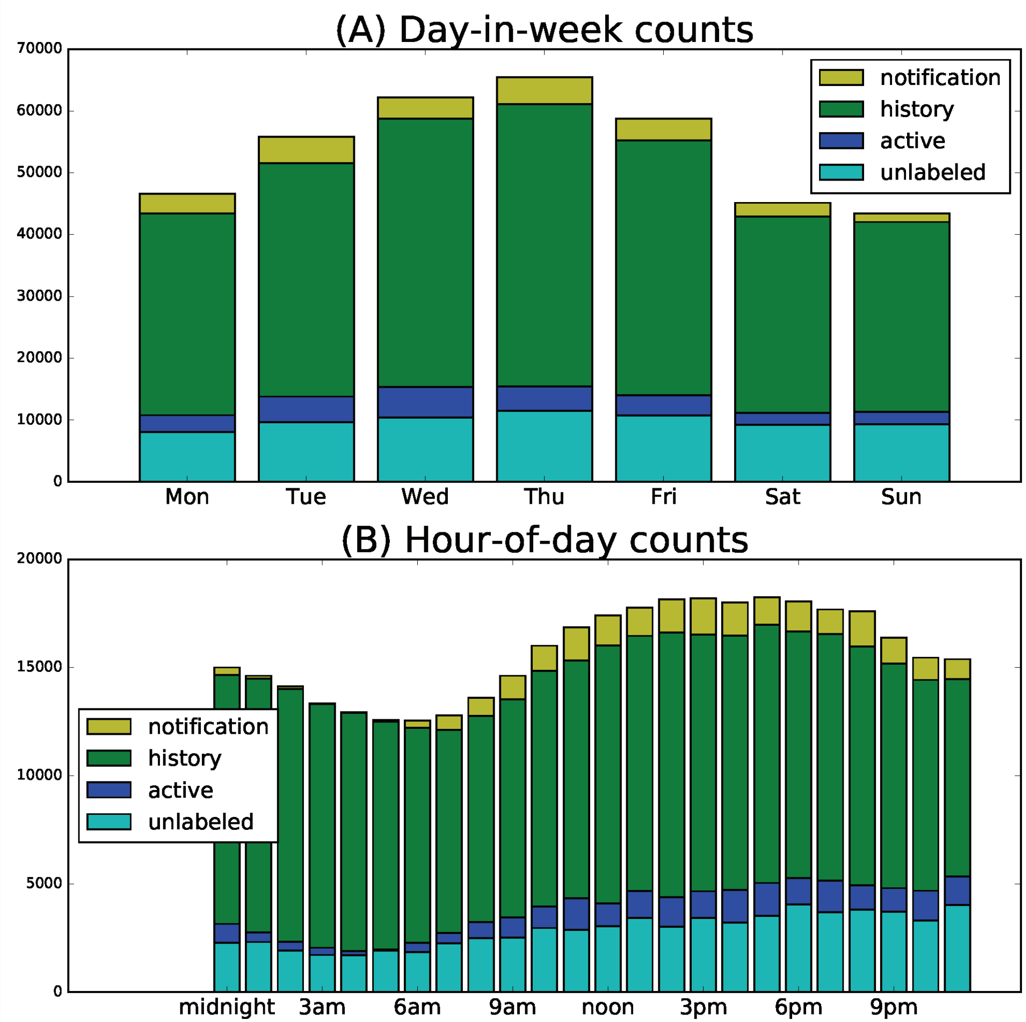
- 60 participants
- ~ 7 days each
- Consent form, compensation ($)

- 308k labeled examples
- Over 50 labels
- ~ 3.8 applied labels per example

- Publicly available: http://extrasensory.ucsd.edu

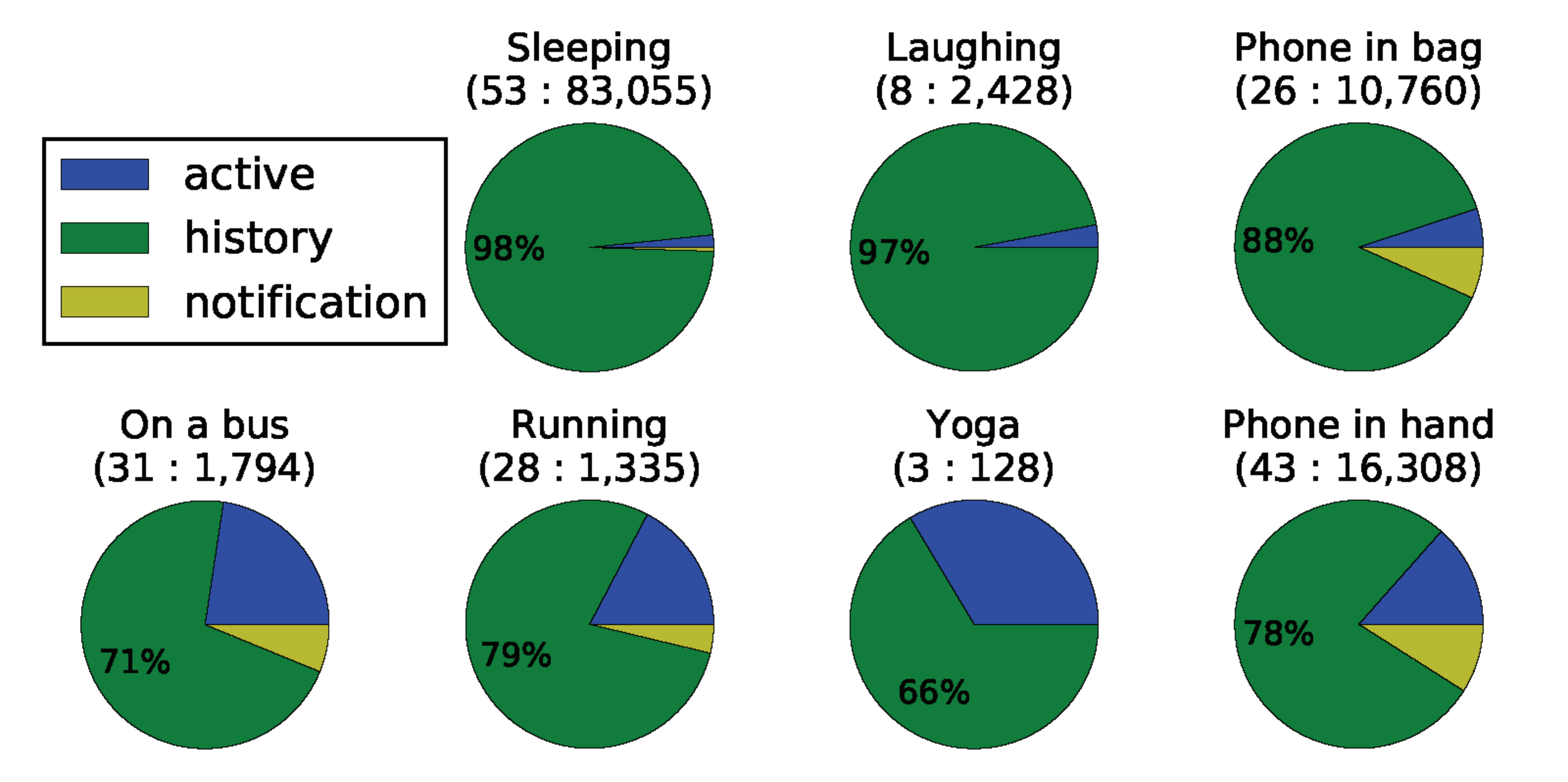
ExtraSensory Dataset:

label-reporting methods



ExtraSensory:

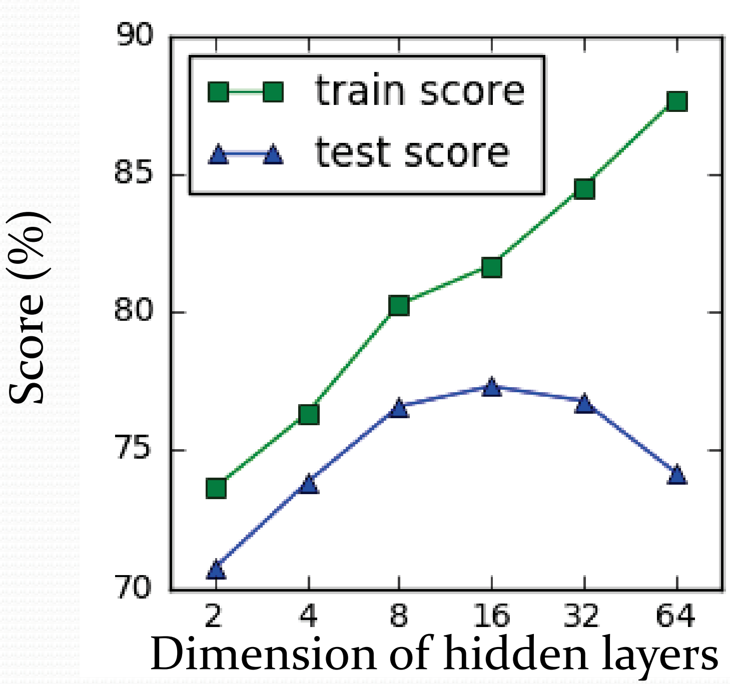
reported context-labels



Recognition evaluation

- Separating train-participants and test-participants
- Fair performance metric (balanced accuracy)

Vaizman, 2017 (December)

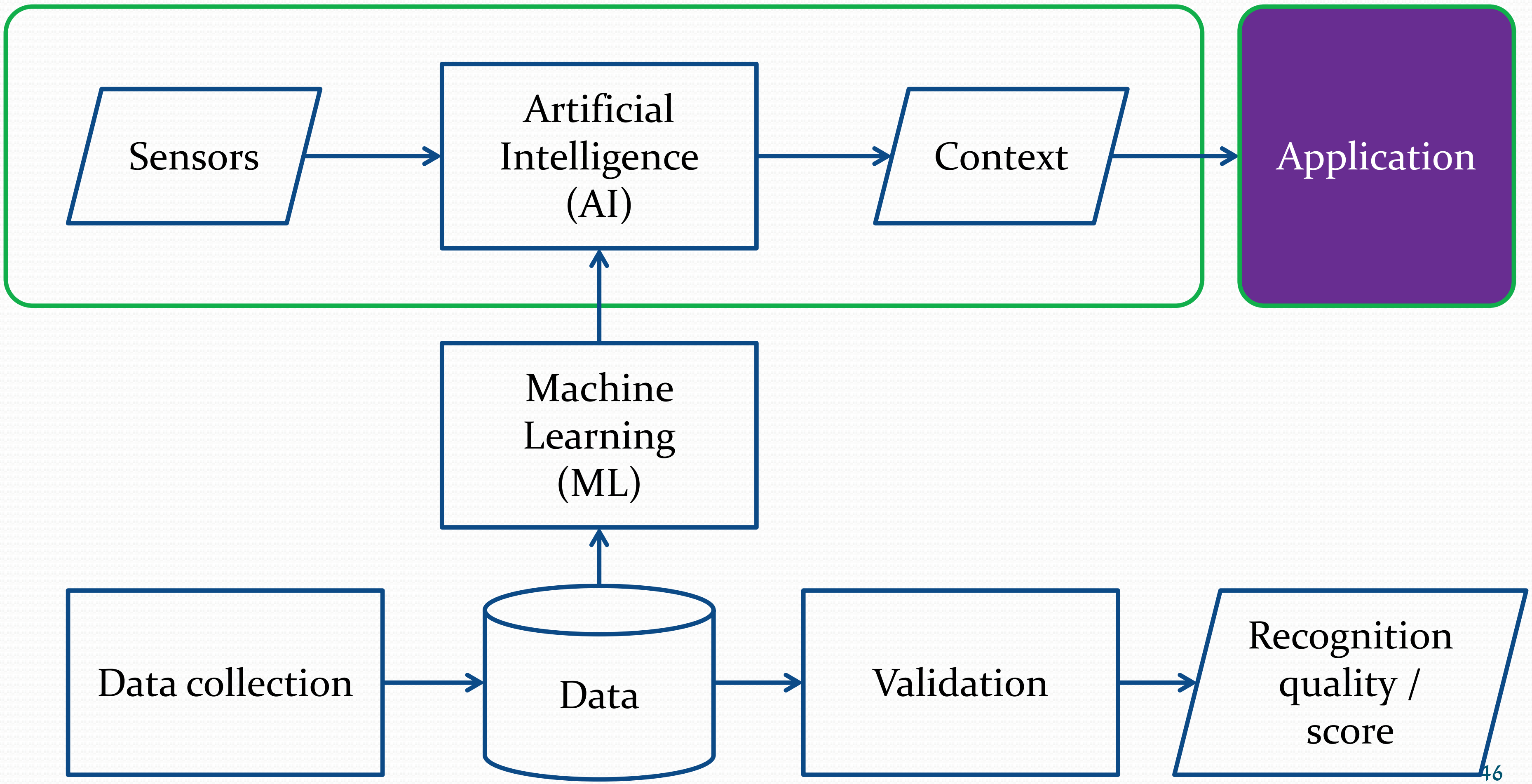


5. Applications

What useful service/benefit can gain with the system?

5. Applications

What useful service/benefit can we gain with the system?



Promote physical activity

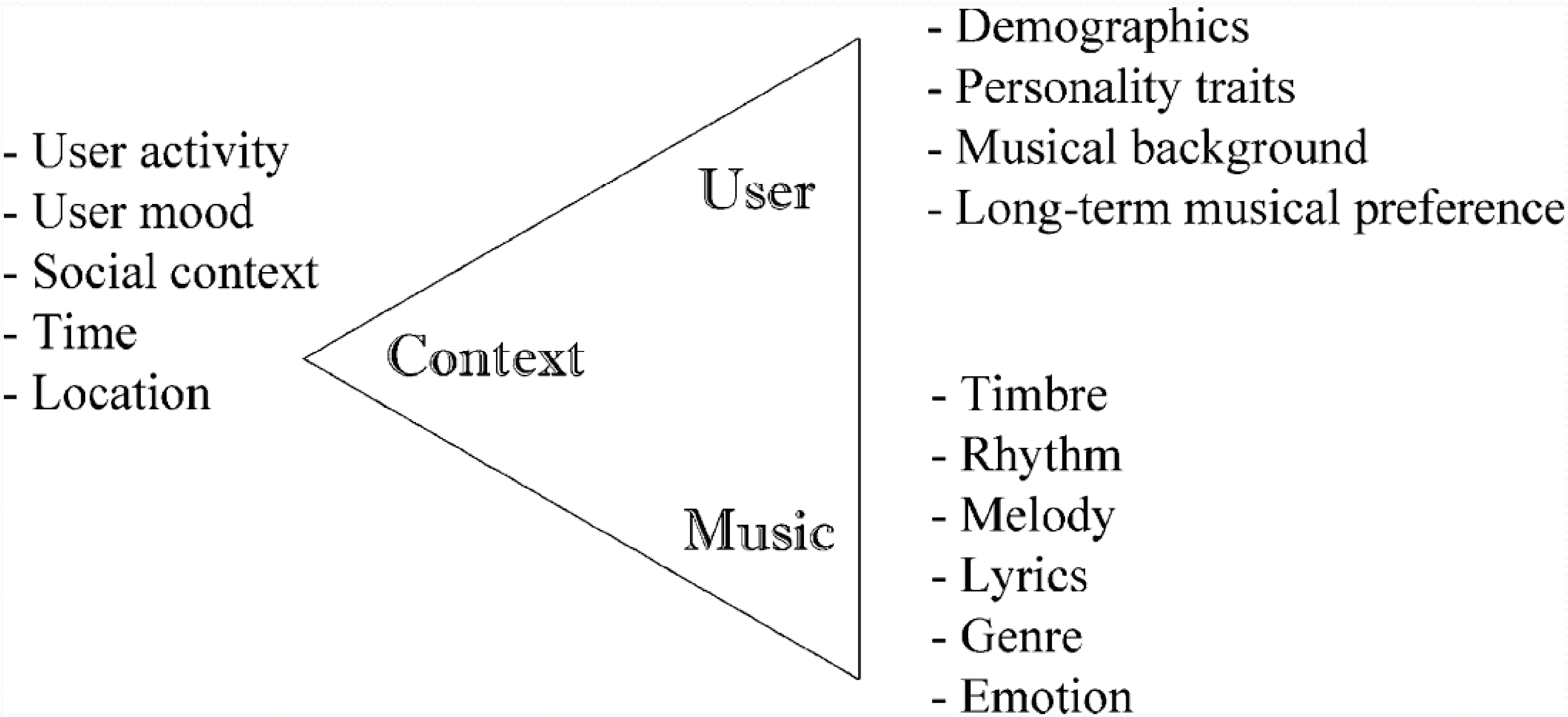
- UbiFit Garden

Consolvo, 2008

Music recommendation

- Personalized, context-aware

Yang, 2015



References

- Bao, L., & Intille, S. (2004). Activity recognition from user-annotated acceleration data. *Pervasive computing*, 1-17.
- Consolvo, S., McDonald, D. W., Toscos, T., Chen, M. Y., Froehlich, J., Harrison, B., ... & Smith, I. (2008, April). Activity sensing in the wild: a field trial of ubifit garden. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 1797-1806). ACM.
- Ellis, K., Kerr, J., Godbole, S., & Lanckriet, G. (2014, September). Multi-sensor physical activity recognition in free-living. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication* (pp. 431-440). ACM.
- Ermes, M., Pärkkä, J., Mäntyjärvi, J., & Korhonen, I. (2008). Detection of daily activities and sports with wearable sensors in controlled and uncontrolled conditions. *IEEE transactions on information technology in biomedicine*, *12*(1), 20-26.
- Ganti, R. K., Srinivasan, S., & Gacic, A. (2010, June). Multisensor fusion in smartphones for lifestyle monitoring. In *Body Sensor Networks (BSN), 2010 International Conference on* (pp. 36-43). IEEE.
- Guiry, J. J., van de Ven, P., & Nelson, J. (2014). Multi-sensor fusion for enhanced contextual awareness of everyday activities with ubiquitous devices. Sensors, 14(3), 5687-5701.
- Hemminki, S., Nurmi, P., & Tarkoma, S. (2013, November). Accelerometer-based transportation mode detection on smartphones. In *Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems* (p. 13). ACM.
- Khan, A. M., Tufail, A., Khattak, A. M., & Laine, T. H. (2014). Activity recognition on smartphones via sensor-fusion and kda-based svms. *International Journal of Distributed Sensor Networks*, 10(5), 503291.
- Pirsiavash, H., & Ramanan, D. (2012, June). Detecting activities of daily living in first-person camera views. In *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on* (pp. 2847-2854). IEEE.
- Tapia, E. M., Intille, S. S., & Larson, K. (2004, March). Activity recognition in the home using simple and ubiquitous sensors. In *Pervasive* (Vol. 4, pp. 158-175).
- Thomaz, E., Zhang, C., Essa, I., & Abowd, G. D. (2015, March). Inferring meal eating activities in real world settings from ambient sounds: A feasibility study. In *Proceedings of the 20th International Conference on Intelligent User Interfaces* (pp. 427-431). ACM.
- Vaizman, Y., Ellis, K., and Lanckriet, G. (2017, October). Recognizing Detailed Human Context In-the-Wild from Smartphones and Smartwatches. *IEEE Pervasive Computing*, vol. 16, no. 4, October-December 2017, pp. 62-74.
- Vaizman, Y., Weibel, N., and Lanckriet, G. (2017, December). Context Recognition In-the-Wild: Unified Model for Multi-Modal Sensors and Multi-Label Classification. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (IMWUT)*, vol. 1, no. 4.
- Yang, Y. H., & Teng, Y. C. (2015). Quantitative study of music listening behavior in a smartphone context. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 5(3), 14.
- Yatani, K., & Truong, K. N. (2012, September). BodyScope: a wearable acoustic sensor for activity recognition. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing* (pp. 341-350). ACM.

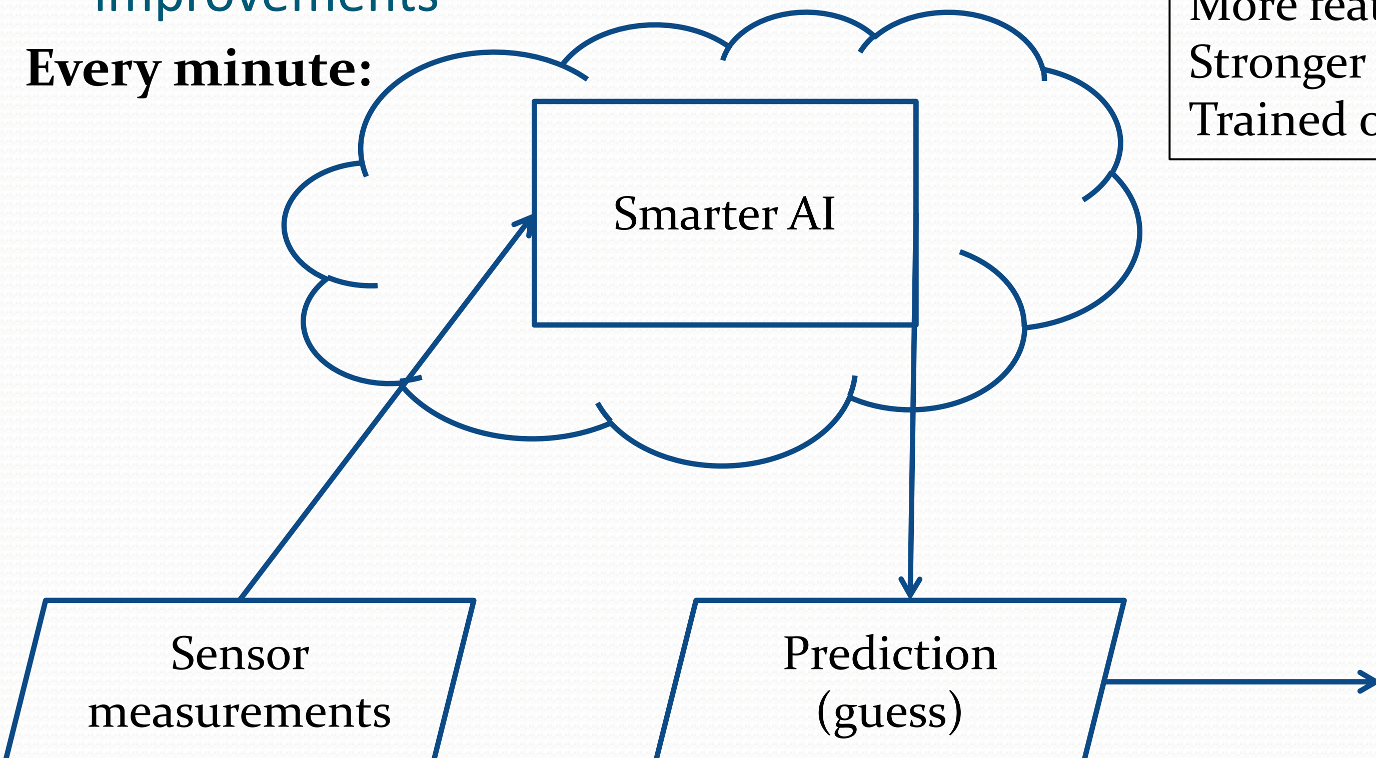
Your class project

- Application using context-recognition
- Provide visualization of user’s behavior
  - Hourly/daily/weekly statistics of contexts
  - Timeline of activities/contexts
  - Contexts presented on a map
  - Your original ideas to make it clear, interesting, fun for user…
- Provide added value
  - Why would anyone want to use it?
- Basic evaluation
  - Simulations, validation with few users
  - After course: exciting applications -> validation with real users

ExtraSensory App:

improvements

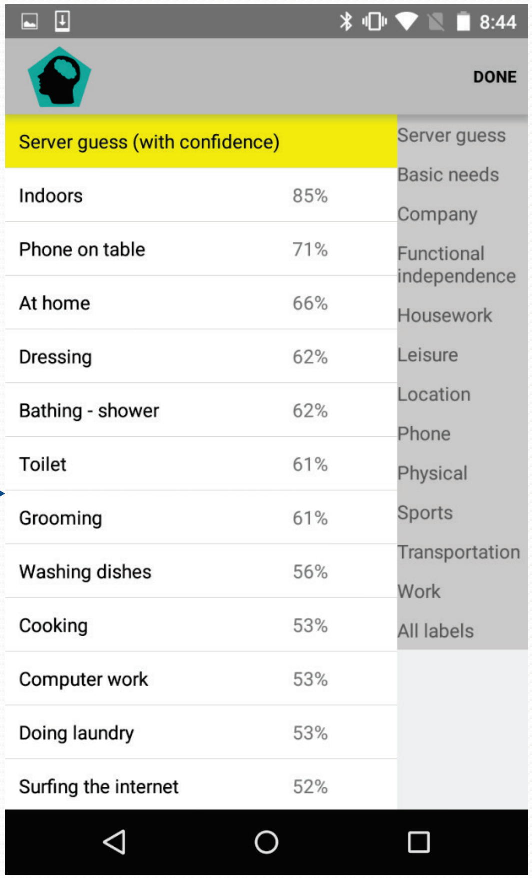
Every minute:



Selected sensors (less battery consumption)

Multi-label
Soft predictions (probabilities)
51 context labels

More features
Stronger classifier model
Trained on ExtraSensory Dataset



ExtraSensory App:

real-time context-recognition

1. Keep it running in background (Android phone)
2. Get per-minute predictions:

Context-label	Probability
Sleeping	0.26
Shower	0.13
Walking	0.45
Sitting	0.67
Computer work	0.87
Indoors	0.90
At a bar	0.24
…	…

3. Get per-minute location lat-long, e.g. (32.88, -117.23)

Possible applications

1. Why would anyone use the application? What is the added value?
2. Who will use the application? When will they use it?
3. How will it work? How will it utilize recognized contexts?

Assume getting per-minute probabilities of these context-labels:

1. Lying down	18. At home	35. Singing
2. Sitting	19. At school	36. Talking
3. Standing	20. At a restaurant	37. Computer work
4. Walking	21. Exercising	38. Eating
5. Running	22. Cooking	39. Toilet
6. Bicycling	23. Shopping	40. Grooming
7. Sleeping	24. Strolling	41. Dressing
8. Lab work	25. Drinking (alcohol)	42. At the gym
9. In class	26. Bathing – shower	43. Stairs – going up
10. In a meeting	27. Cleaning	44. Stairs – going down
11. At work	28. Doing laundry	45. Elevator
12. Indoors	29. Washing dishes	46. Phone in pocket
13. Outside	30. Watching TV	47. Phone in hand
14. In a car	31. Surfing the internet	48. Phone in bag
15. On a bus	32. At a party	49. Phone on table
16. Drive – I'm the driver	33. At a bar	50. With co-workers
17. Driver – I'm the passenger	34. At the beach	51. With friends

Next step

- Piazza – discuss project ideas
- Form teams (3 undergrads, 3 grads) – by Sunday Oct 15
- Submit project proposal – by Sunday Oct 22

- This Thursday (Oct 12) at 11:00am – technical session about using the ExtraSensory App.