UNI KASSEL VERSITÄT

Chair for Communication Technology

Prof. Dr.-Ing. Klaus David

ComTec

Communication Technologies 2 (CT2)
Machine Learning: Applications and Algorithms

Activity Recognition: Evaluation metrics

Michel Morold

Lecture in WS 2018 / 2019
01.11.2018

© ComTec 2018

Outline

- Activity Recognition
- Instance-based Activity Recognition
- Event-based Evaluation
- Summary

© ComTec 2018

ACTIVITY RECOGNITION (AR)

© ComTec 2018

Activity Recognition (AR)

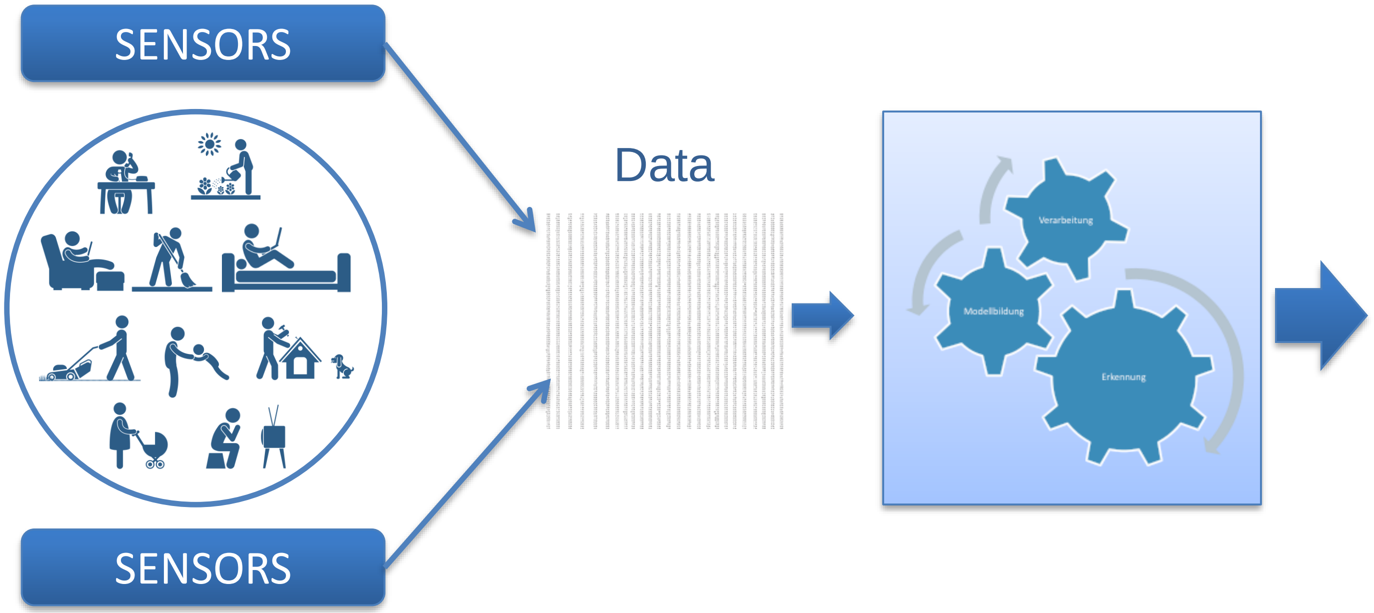
…to recognize the actions and intentions of a user based on a group of observations

Daily Life Activities

© ComTec 2018

Activity Recognition (AR)

ComTec

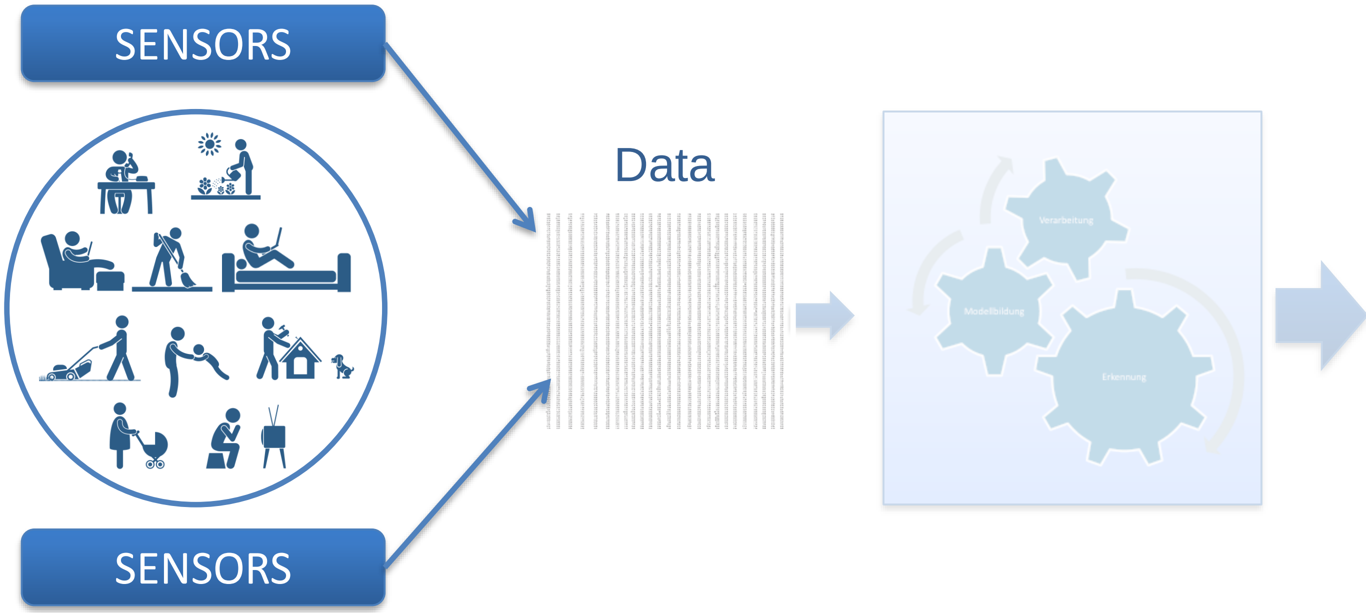


- Laying
- Walking
- Standing
- Cooking
- Going up stairs
- Driving
- etc...

© ComTec 2018

Activity Recognition (AR)

ComTec



- Laying
- Walking
- Standing
- Cooking
- Going up stairs
- Driving
- Etc...

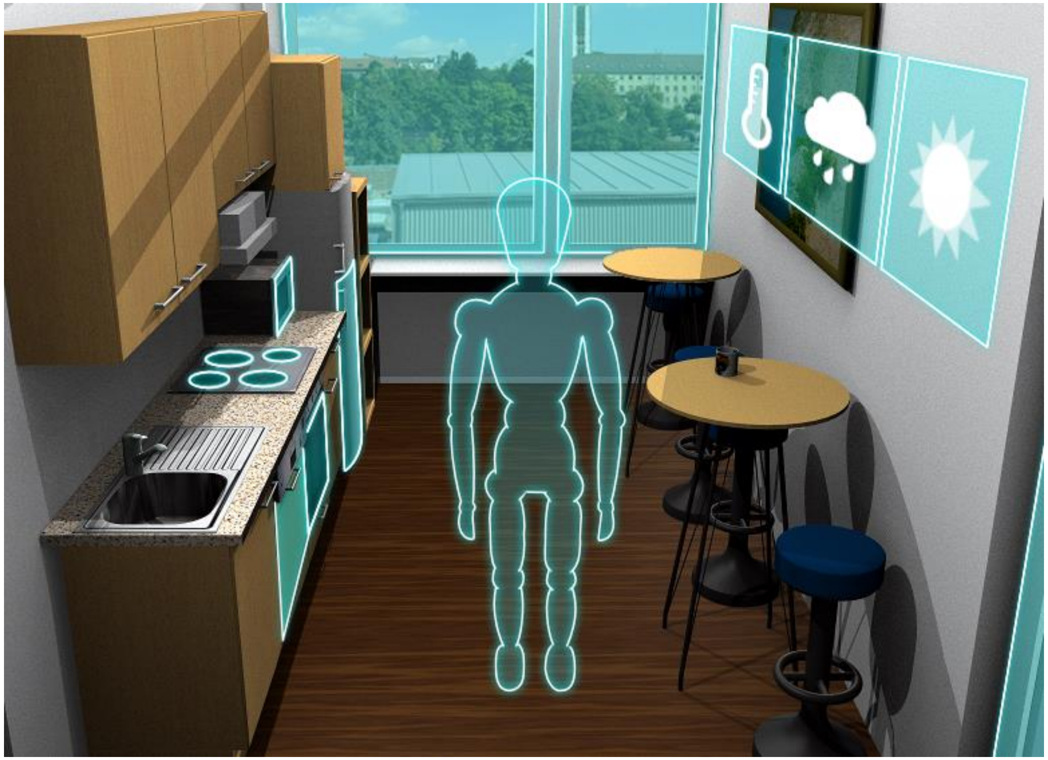
© ComTec 2018

Sensor-based Activity Recognition

ComTec

Dense Sensing

– Attached to objects to monitor human activities through user-object interactions
– Suitable for activities that involve a number of objects within an environment



[ComTec]

© ComTec 2018

Sensor-based Activity Recognition

ComTec

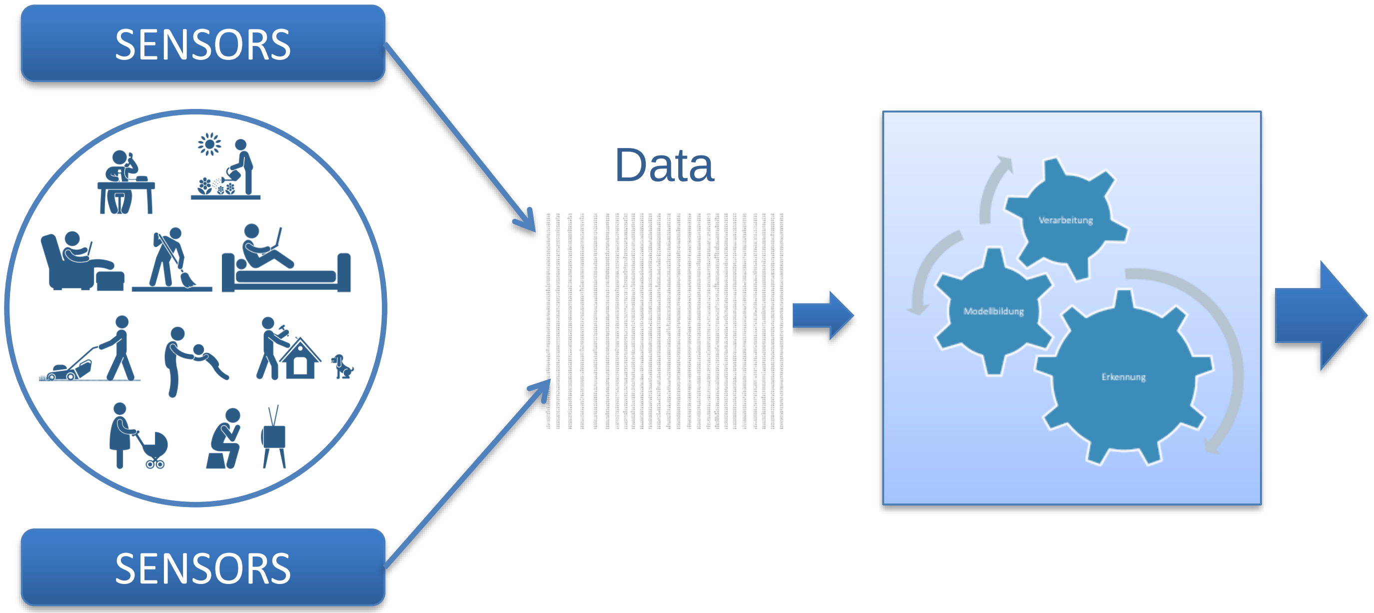
Wearable Sensing

- Can be positioned directly or indirectly on the body
- Sensors be embedded into, clothes, mobile devices, etc.
- Position, pulse, skin temperatures
- Mainly physical activities like walking, running, etc.

[1]

© ComTec 2018

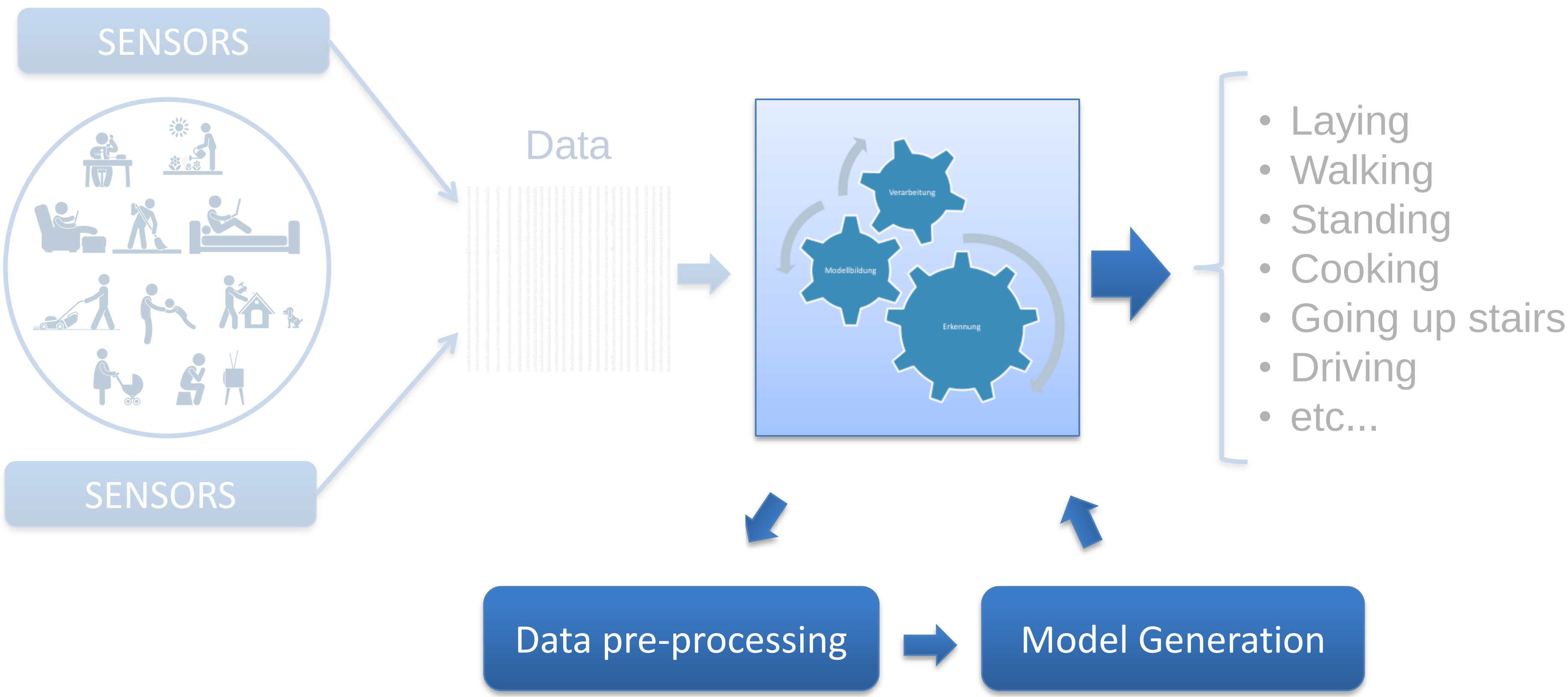
Activity Recognition Process



- Laying
- Walking
- Standing
- Cooking
- Going up stairs
- Driving
- etc...

© ComTec 2018

Activity Recognition Process



- Segmentation
- Feature extraction

- Machine Learning

© ComTec 2018

Machine learning

ComTec

Supervised

- Bayesian statistics
- Decision trees
- Artificial neural network
- Support vector machines
- Hidden Markov models
- Etc…

Unsupervised

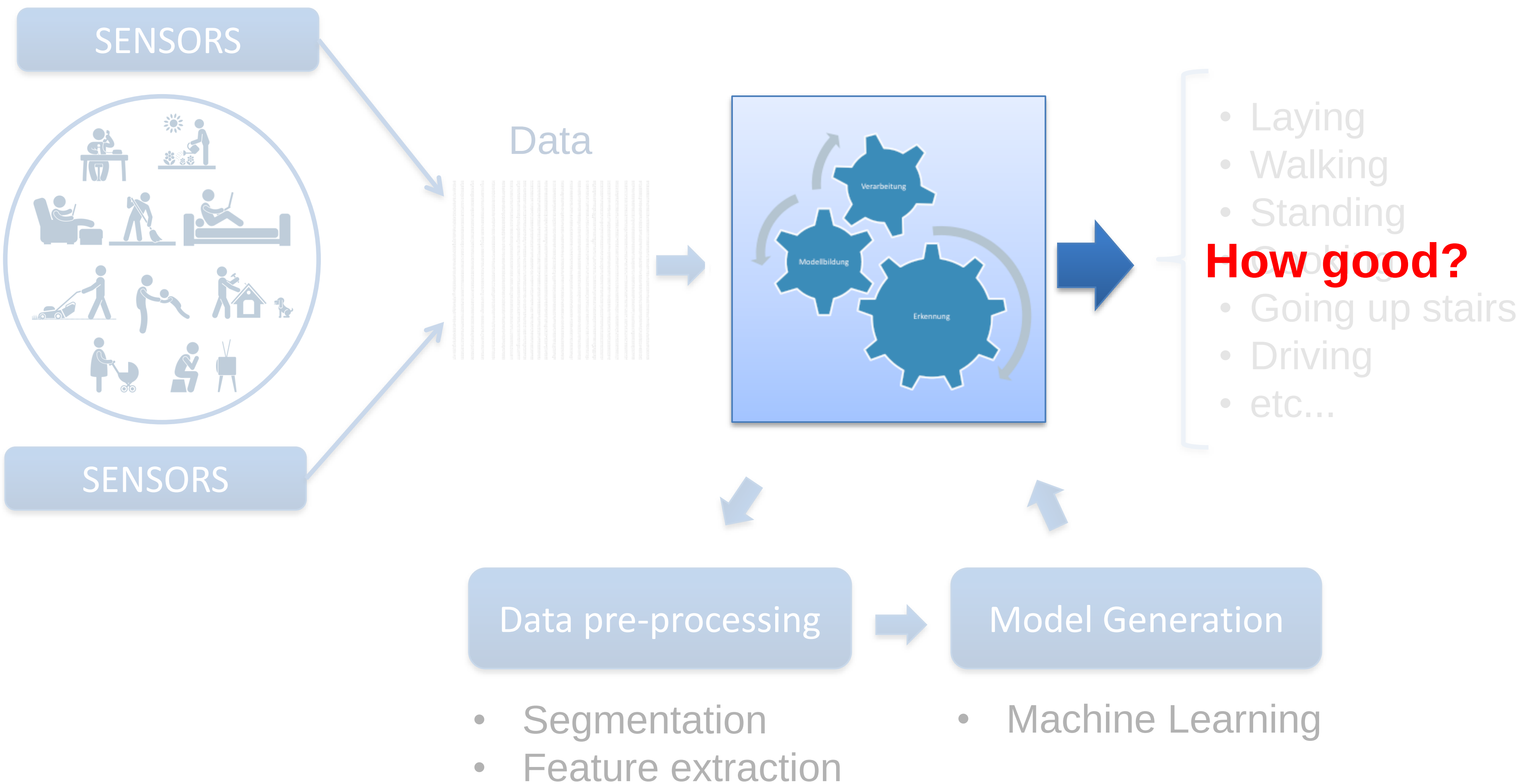
- Data clustering
- Self-organizing map
- Artificial neural network
- Expectation-maximization algorithm
- Etc…

Reinforcement

- Monte Carlo Method
- Q-learning
- Temporal difference learning
- Learning Automata
- Etc…

© ComTec 2018

Activity Recognition Process



INSTANCE-BASED ACTIVITY RECOGNITION

Instance-based AR

ComTec

- An instance is a vector that contains all features calculated within a segment
- For supervised learning the instance is labeled with an activity

time

	Ground Truth	Recognized by the model
1	step down	step down
2	walking	step down
3	walking	walking
4	sitting	sitting
…	…	…

© ComTec 2018

Example – 3 activities / 300 instances

Activity	Total count of Instances
step down	100
walking	100
sitting	100
∑	300

Example – 3 activities / 300 instances

ComTec

Activity	Total count of Instances
step down	100
walking	100
sitting	100
∑	300

time

Ground Truth	Recognized
step down	step down
walking	step down
walking	walking
sitting	sitting
…	…

Example – 3 activities / 300 instances

Activity	Total count of Instances	Correct recognized Instances
step down	100	100
walking	100	95
sitting	100	90
∑	300	285

time

Ground Truth	Recognized
step down	step down
walking	step down
walking	walking
sitting	sitting
…	…

© ComTec 2018

Example – 3 activities / 300 instances Confusion Matrix

Recognition

step down	walking	sitting		Ground truth
100	0	0	**step down**	Σ 100
5	**95**	0	**walking**	Σ 100
7	3	**90**	**sitting**	Σ 100
Σ 112	Σ 98	Σ 90		Σ 300

© ComTec 2018

What is recognition accuracy over all classes?

ComTec

Recognition

step down	walking	sitting		Ground truth
100	0	0	**step down**	Σ 100
5	**95**	0	**walking**	Σ 100
7	3	**90**	**sitting**	Σ 100
Σ 112	Σ 98	Σ 90		Σ 300

- recognition performance averaged over all activities

$$classification\ accuracy = \frac{correct\ recognized\ instances}{total\ count\ of\ instances} = \frac{100 + 95 + 90}{300} = 95\%$$

© ComTec 2018

What is recognition performance of activity “step down”?

- Recognition performance as binary problem e.g. „step down“ or „not step down“
- represented in a tuple of two elements “AB”
  - B: can be Positive (P) or Negative (N)
    - Positive (P): classifier recognizes “step down”
    - Negative (N): classifier detects “not step down”, e.g. walking was recognized instead
  - A: can be True (T) of False (F)
    - True (T): the recognition is correct
    - False (F): the recognition is not correct
- ”AB“ results in four combinations: TP, FP, TN, FN

Tuple combinations

Recognition

Ground truth	step down	not step down
step down	TP	FN
not step down	FP	TN

- *True Positives* (TP):
  - Recognition step down (positive) and this classification is correct (true)
- *False Positives* (FP):
  - Recognition step down (positive) and this classification is incorrect (false)
- *True Negatives* (TN):
  - Recognition not step down (negative) and this classification is correct (true)
- *False Negatives* (FN):
  - Recognition not step down (negative) and this classification is incorrect (false)

- origin in Information Retrieval

© ComTec 2018

What is recognition performance of activity “step down”?

Recognition

step down	not step down	not step down		Ground truth
100	0	0	**step down**	Σ 100
5	95	0	**not step down**	Σ 100
7	3	90	**not step down**	Σ 100
Σ 112	Σ 98	Σ 90		Σ 300

- binary classification problem focused on activity “step down”
  - **TP** = ?

What is recognition performance of activity “step down”?

Recognition

step down	not step down	not step down		Ground truth
100	0	0	step down	Σ 100
5	95	0	not step down	Σ 100
7	3	90	not step down	Σ 100
Σ 112	Σ 98	Σ 90		Σ 300

- binary classification problem focused on activity “step down”
  - **TP** = 100
  - **FP** = ?

© ComTec 2018

What is recognition performance of activity “step down”?

ComTec

Recognition

step down	not step down	not step down		Ground truth
100	0	0	**step down**	Σ 100
5	95	0	**not step down**	Σ 100
7	3	90	**not step down**	Σ 100
Σ 112	Σ 98	Σ 90		Σ 300

- binary classification problem focused on activity “step down”
  - **TP** = 100
  - **FP** = 5+7 = 12
  - **TN** = 95+0+3+90 = 188
  - **FN** = ?

What is recognition performance of activity “step down”?

ComTec

Recognition

step down	not step down	not step down	Ground truth	
100	0	0	step down	Σ 100
5	95	0	not step down	Σ 100
7	3	90	not step down	Σ 100
Σ 112	Σ 98	Σ 90		Σ 300

- binary classification problem focused on activity “step down”
  - **TP** = 100
  - **FP** = 5+7 = 12
  - **TN** = 95+0+3+90 = 188
  - **FN** = 0+0 = 0

© ComTec 2018

Precision

The precision, often referred to *positive predictive value*, represents the ratio of correctly classified (truly) “step down” instances to the total number of instances recognized as “step down”.

Therefor the precision expresses the “step down” detection accuracy.

↓ **precision**

Recognition

step down	**not step down**	**not step down**	Ground truth	
100	0	0	**step down**	Σ 100
5	95	0	**not step down**	Σ 100
7	3	90	**not step down**	Σ 100
Σ 112	Σ 98	Σ 90		Σ 300

$$Precision = \frac{TP}{(TP + FP)}$$

step down (not step down e.g. walking or sitting)

precision = 100 / (100+12) = 89.3%

© ComTec 2018

Recall

ComTec

The recall, also called *true positive rate* or *hit rate* represents the ratio of correctly classified “step down” instances to the total number of “step down” instances, i.e. how many of the truly “step downs” were correctly classified.

The recall represents the usability of the detection as a low usability leads to high rate of undetected “step downs”.

↓ precision

Recognition

	step down	**not step down**	**not step down**		
step down	100	0	0	Σ 100	← **recall**
not step down	5	95	0	Σ 100	
not step down	7	3	90	Σ 100	
	Σ 112	Σ 98	Σ 90	Σ 300	

Ground truth

$$Recall = \frac{TP}{(TP + FN)}$$

step down (not step down e.g. walking or sitting)

recall = 100 / (100+0) = 100.0%

© ComTec 2018

F-measure

ComTec

$$F = \frac{2 * (Precision * Recall)}{(Precision + Recall)}$$

↓ precision
Recognition

	step down	not step down	not step down		Ground truth	← recall
	100	0	0	step down		Σ 100
	5	95	0	not step down		Σ 100
	7	3	90	not step down		Σ 100
	Σ 112	Σ 98	Σ 90			Σ 300

step down (not step down e.g. walking or sitting)

precision	= 100 / (100+12) =	89.3%
recall	= 100 / (100+0) =	100.0%
F-measure	= 2 * (89.3% * 100.0%) / (89.3% + 100.0%) =	94.3%

© ComTec 2018

Metrics summary

↓ precision
Recognition

step down	not step down	not step down		
100	0	0	step down	Σ 100
5	95	0	not step down	Σ 100
7	3	90	not step down	Σ 100
Σ 112	Σ 98	Σ 90		Σ 300

Ground truth

← recall

step down (not step down e.g. walking or sitting)

precision = 100 / (100+12) = 89.3%
recall = 100 / (100+0) = 100.0%
F-measure = 2 * (89.3% * 100.0%) / (89.3% + 100.0%) = 94.3%

walking (not walking e.g. step down or sitting)

precision = 95 / (95+3) = 96.9%
recall = 95 / (95+5) = 95.0%
F-measure = 2 * (96.9% * 95.0%) / (96.9% + 95.0%) = 96.0%

sitting (not sitting e.g. step down or walking)

precision = 90 / (90+0) = 100.0%
recall = 90 / (90+10) = 90.0%
F-measure = 2 * (100.0% * 90.0%) / (100.0% + 90.0%) = 94.7%

© ComTec 2018

Instance-based evaluation

- An instance (or frame) is a fixed-length, fixed-rate unit of time
- Instance by instance comparison between prediction (i.e. recognition) and ground truth (i.e. labels)
- Does not consider specific error types of continuous activity recognition

© ComTec 2018

Instance-based evaluation

- Challenges detecting real world activities
  - Activities can have variable durations and ambiguous start and stop times, so that activities are being detected before or after they actually occur
  - Single events being fragmented into multiple smaller events of the same activity
  - Merging of several real events into a single recognized event

➡ Event-based Evaluation (Ward et al. [2])

Event-based Evaluation

ComTec

- An *event* is a variable duration sequence of positive instances within a continuous time series (i.e. a sequence of the same instances)

Instances
- step down
- walking
- walking
- walking
- sitting
- walking
- ...

Events
- step down
- walking
- sitting
- walking
- ...

Event-based Evaluation

- A *segment* is a contiguous sequence of instances with a variable duration, in which neither prediction nor the ground truth labels changes

Event-based Evaluation
Type of Errors

- Ground truth Errors
  - Fragmentation
  - Deletion
  - Underfill

- Recognition Errors
  - Merge
  - Insertion
  - Overfill

© ComTec 2018

Event-based Evaluation - Ground truth Errors

- Fragmentation (F): A ground truth event that contains more than one matching segment

Ground truth	Recognition
step down	step down
walking	walking
walking	sitting
walking	walking
sitting	sitting
walking	walking
...	...

Event-based Evaluation - Ground truth Errors

- Deletion (D): A ground truth event that contains no matching segments

Ground truth	Recognition
step down	step down
walking	walking
walking	walking
walking	walking
sitting	walking
walking	walking
...	...

© ComTec 2018

Event-based Evaluation - Ground truth Errors

ComTec

- Underfill (U): A ground truth event not completely covered by recognition

Ground truth	Recognition
step down	step down
walking	walking
walking	walking
walking	sitting
sitting	sitting
walking	walking
…	…

Event-based Evaluation - Recognition Errors

ComTec

- Merge (M): A recognized event that contains more than one matching segment

Ground truth	Recognition
step down	step down
walking	walking
walking	walking
walking	walking
sitting	walking
walking	walking
…	…

© ComTec 2018

Event-based Evaluation - Recognition Errors

ComTec

- Insertion (I): A recognized event that contains no matching segments

Ground truth	Recognition
step down	step down
walking	walking
walking	sitting
walking	walking
sitting	sitting
…	…
…	…

© ComTec 2018

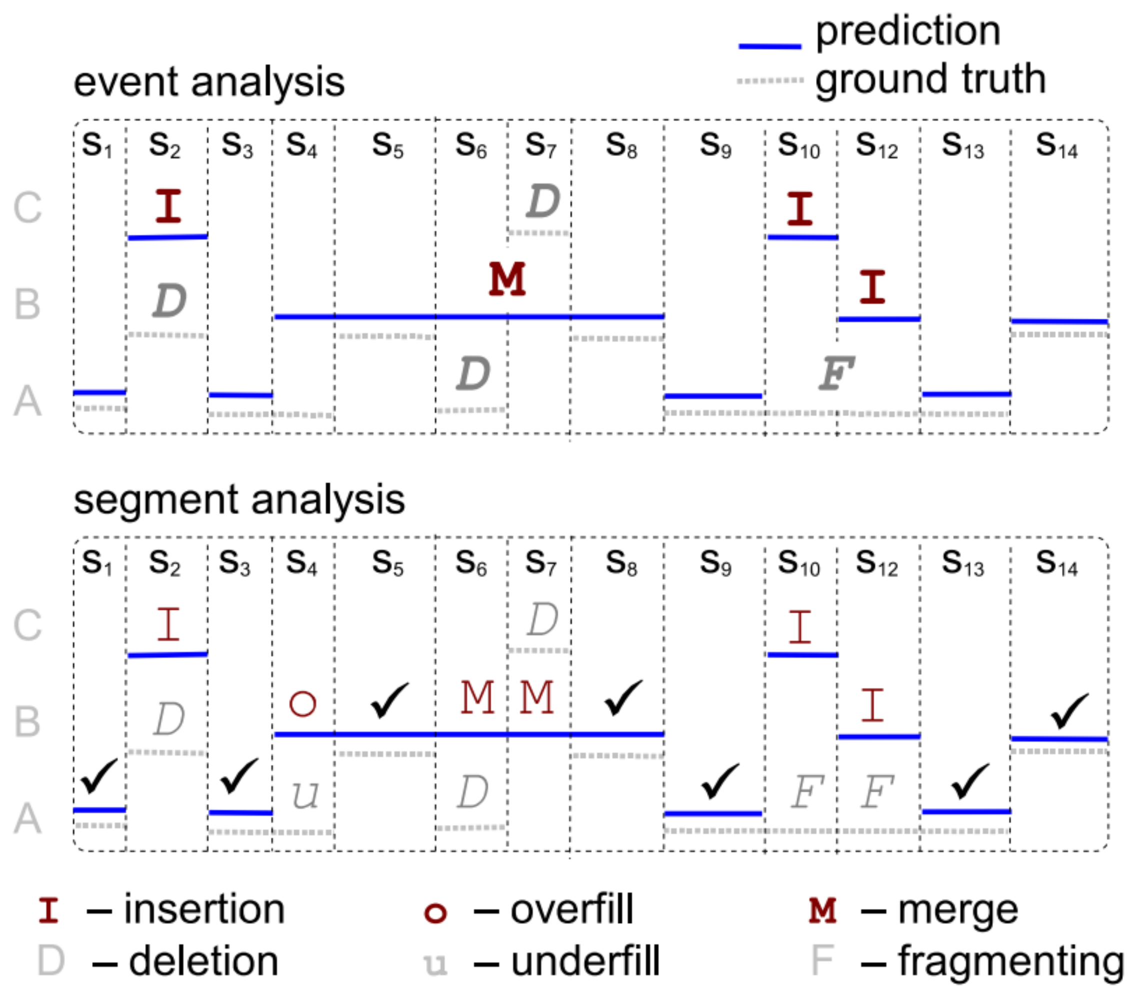
Event-based Evaluation - Recognition Errors

- Overfill (O): A recognized event which spills over its ground truth boundary

Ground truth	Recognition
step down	step down
walking	walking
walking	walking
walking	walking
sitting	walking
sitting	sitting
...	...

© ComTec 2018

Example

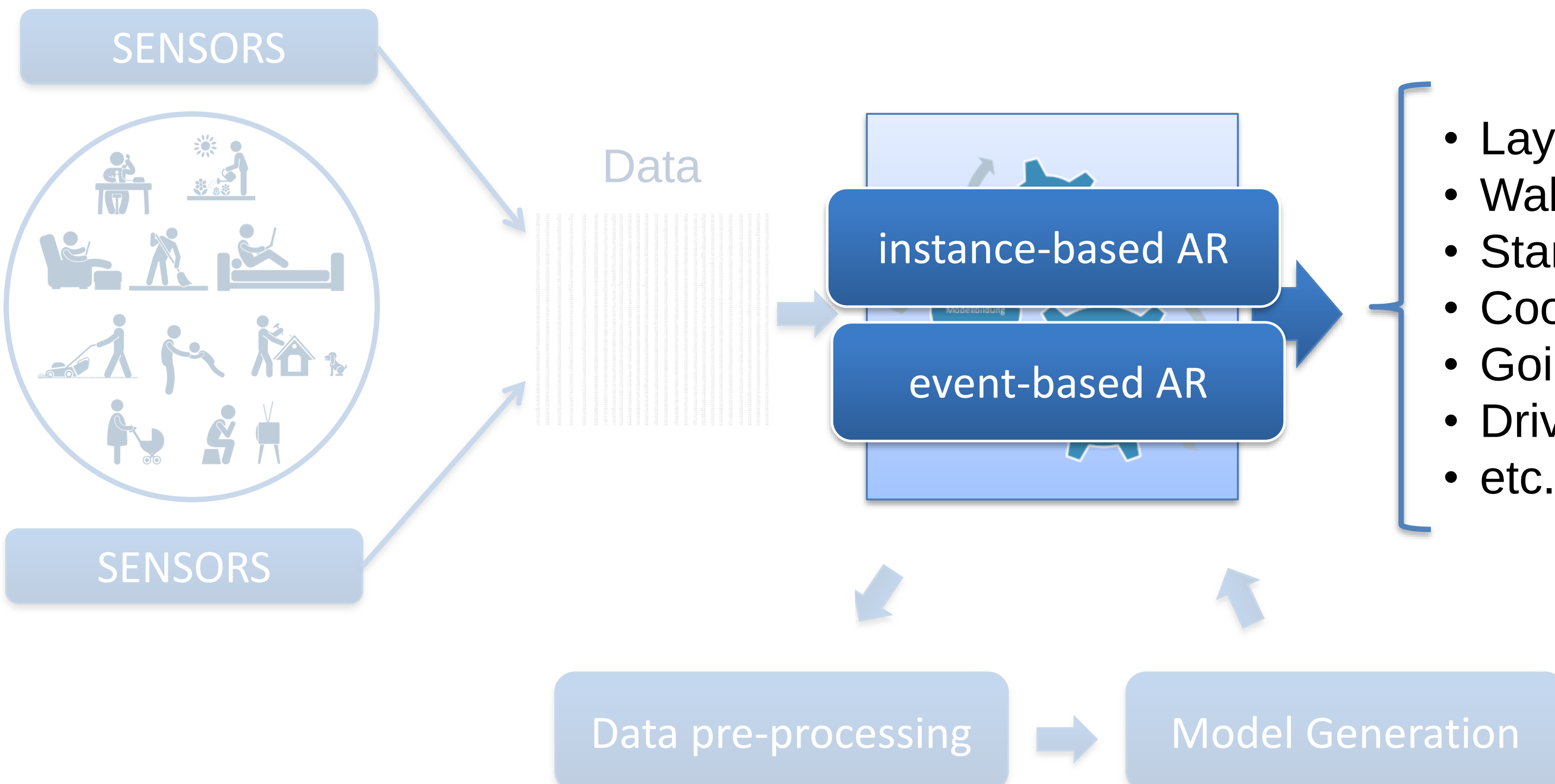


[4]

© ComTec 2018

Activity Recognition Process

ComTec



- Laying
- Walking
- Standing
- Cooking
- Going up stairs
- Driving
- etc...

- Segmentation
- Feature extraction

- Machine Learning

Summary Questions

1. What is the purpose of activity recognition?
2. What is the difference between online and offline activity recognition?
3. What is the purpose of performance metrics?
4. Name one metric that show a models recognition performance for one activity.
5. What does the event-based evaluation help?
6. Name some errors in the event-based evaluation.

References

[1] L. Bao and S. S. Intille, “Activity Recognition from User-Annotated Acceleration Data,” in Lecture Notes in Computer Science, Pervasive Computing, T. Kanade, J. Kittler, J. M. Kleinberg, F. Mattern, J. C. Mitchell, O. Nierstrasz, C. Pandu Rangan, B. Steffen, D. Terzopoulos, D. Tygar, M. Y. Vardi, and A. Ferscha, Eds, Berlin, Heidelberg: Springer Berlin Heidelberg, 2004, pp. 1–17.

[2] J. A. Ward, P. Lukowicz, and G. Tröster, “Evaluating Performance in Continuous Context Recognition Using Event-driven Error Characterisation,” in Proceedings of the Second International Conference on Location- and Context-Awareness, Berlin, Heidelberg, 2006, pp. 239–255.

[3] J. A. Ward, P. Lukowicz, and H. W. Gellersen, “Performance Metrics for Activity Recognition,” ACM Trans. Intell. Syst. Technol., vol. 2, no. 1, pp. 1–23, Jan. 2011.

[4] J. A. Ward, P. Lukowicz, and G. Tröster, “A categorisation of performance errors in continuous context recognition,” in Proc. IEEE Int’l Symp. on Wearable Comp, Workshop for On-Body Sensing, 2005.

© ComTec 2018