Machine Learning and Pervasive Computing

Stephan Sigg

Georg-August-University Goettingen, Computer Networks

22.10.2014

Overview and Structure

22.10.2014	Organisation
22.10.3014	Introduction (Def.: Machine learning, Supervised/Unsupervised, Examples)
29.10.2014	Machine Learning Basics (Toolchain, Features, Metrics, Rule-based)
05.11.2014	A simple Supervised learning algorithm
12.11.2014	Excursion: Avoiding local optima with random search
19.11.2014	–
26.11.2014	Bayesian and decision-tree learner
03.12.2014	–
10.12.2014	Non-parametric methods
17.12.2014	Higher dimensional data (SVM, ANN, SOM)
07.01.2015	Classification (Single class, multi-class)
14.01.2015	Unsupervised learning
21.01.2015	Dimensionality reduction (Motivation, PCA)
28.01.2015	Anomaly detection
04.02.2015	Online learning and Recommender systems

Outline

What is machine learning ?

What is machine learning ?

Wikipedia

Machine learning is a subfield of computer science and artificial intelligence that deals with the construction and study of systems that can learn from data, rather than follow only explicitly programmed instructions.

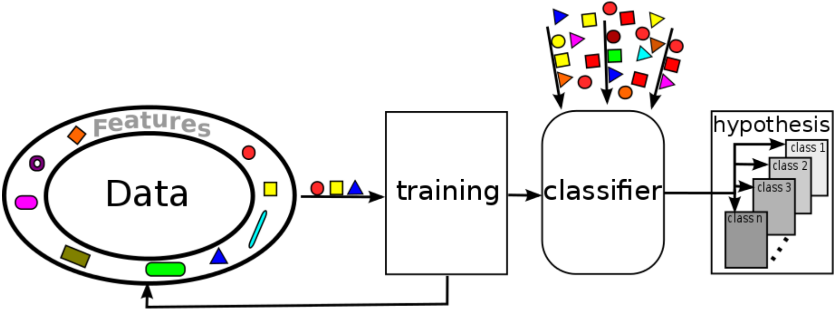
What is machine learning ?

Wikipedia

Machine learning is a subfield of computer science and artificial intelligence that deals with the construction and study of systems that can learn from data, rather than follow only explicitly programmed instructions.

Tom M. Mitchell, The Discipline of Machine Learning, 2006

Machine Learning seeks to answer the question how we can we build computer systems that automatically improve with experience, and what are the fundamental laws that govern all learning processes'


Features
Data
training
classifier
hypothesis
class 1
class 2
class 3
class n

Artificial intelligence

Anomaly detection

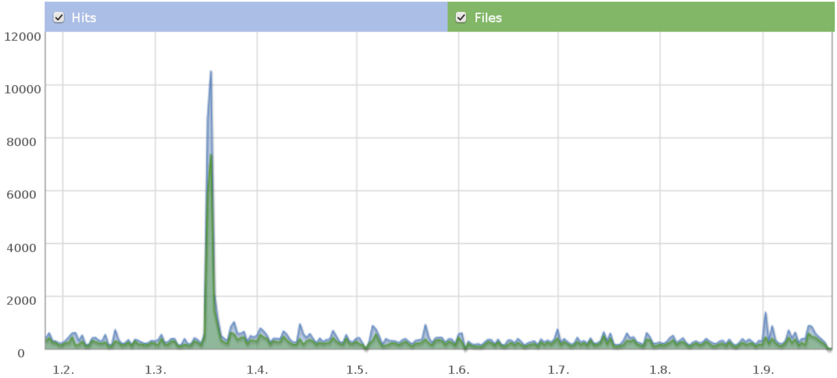


Image segmentation

Image segmentation

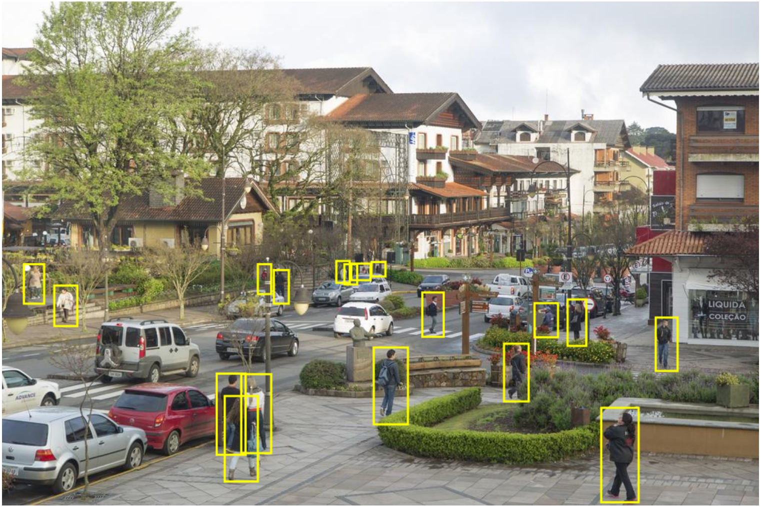


Image segmentation

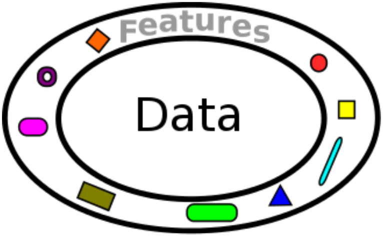
Image segmentation

Image segmentation

Recommender systems

Here: http://www.movielens.org

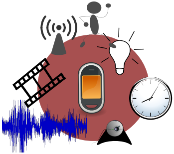
Predictions for you	Your Ratings	Movie Information
★★★★½	Not seen	**Black Christmas (1974)** DVD VHS info | imdb | add tag Comedy, Horror, Mystery, Thriller
★★★★½	Not seen	**Blood of Heroes, The (1988)** DVD info | imdb Action, Sci-Fi
	[add tag] Popular tags:	australia, fighting, delroy lindo
★★★★½	Not seen	**Sisters (1973)** DVD VHS info | imdb | add tag Horror, Thriller
★★★★½	Not seen	**Unsane (Tenebre) (1982)** DVD VHS info | imdb Crime, Horror, Mystery, Thriller - Italian, English
	[add tag] Popular tags:	Giallo, Italian
★★★★½	Not seen	**Collector, The (1965)** DVD VHS info | imdb | add tag Drama, Horror, Thriller
★★★★½	Not seen	**Howling, The (1980)** DVD VHS info | imdb Horror, Mystery
	[add tag] Popular tags:	werewolf
★★★★½	Not seen	**Dawn of the Dead (2004)** DVD VHS info | imdb Action, Drama, Horror, Thriller
	[add tag] Popular tags:	mmm... brains..., Zombie Movie, bmovie
★★★★½	Not seen	**Dark Water (Honogurai mizu no soko kara) (2002)** DVD VHS info | imdb Horror, Mystery, Thriller - Japanese


Features
Data

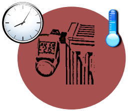
Sensors and sensor classes

- We are surrounded by a multitude of sensors

- Sensor readings utilised for
  - Information provisioning
  - Situation classification
  - Authentication
  - Cryptography



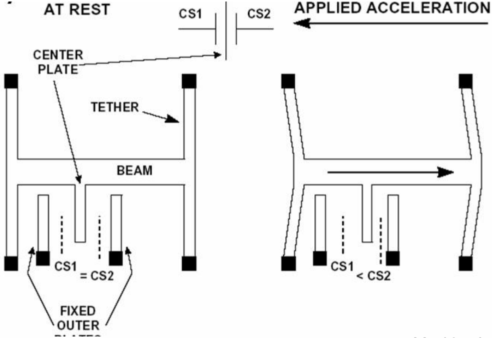
Time
RF
Temperature
Light
Audio
Video
Smoke
Fingerprint
...



Sensors and sensor classes

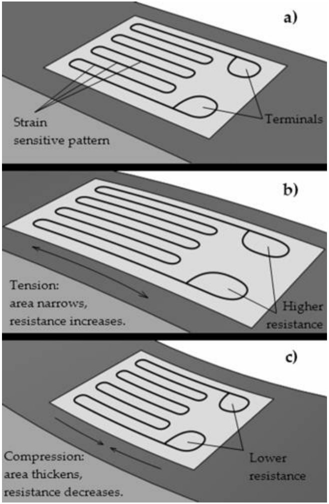
Sensors and sensor classes

- MEMS acceleration sensors
  - E.g. Analogue Devices ADXL
  - Low energy consumption, small, cheap, medium precision
  - MEMS = Micro-mechanical System: Mechanic in Silicon (Silizium)
  - Here: Comparison of capacity CS1 and CS2 leads to acceleration



Sensors and sensor classes

- Pressure sensors
  - Z.B. IEE about 3-10 Euro
  - Very imprecise

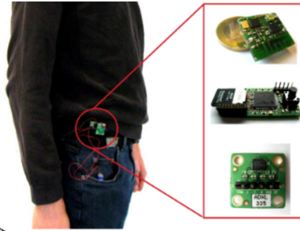


Sensors and sensor classes

- Output of sensors has to be interpreted typically
  - Raw electrical signals
  - Interpretation of signals as electric values
  - Binary or Real valued representation
  - Further identification of features
  - Feature extraction
  - Interpretation of features and classification

Data sampling

- Record <u>sufficient</u> training data
  - Annotated! (Ground-truth)
  - Multiple subjects
  - Various environmental conditions (time of day, weather, ...)

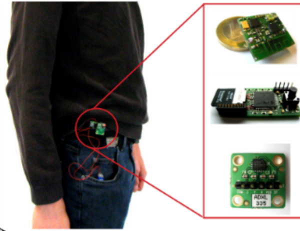


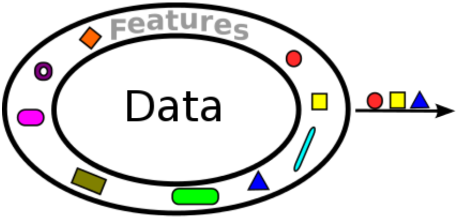
Data sampling

- Record <u>sufficient</u> training data
  - Annotated! (Ground-truth)
  - Multiple subjects
  - Various environmental conditions (time of day, weather, ...)

Example

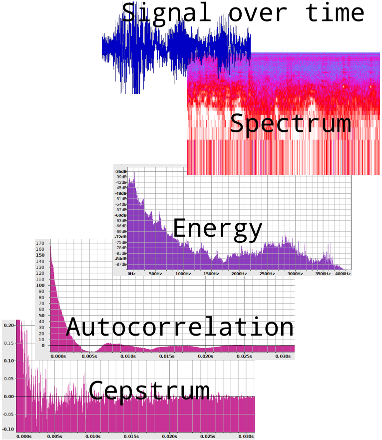
- Electric supply data over 15 years covers 5000 days but only 15 christmas days
- Especially critical events like accidents (e.g. plane, car, earthquake) are scarce




Features
Data

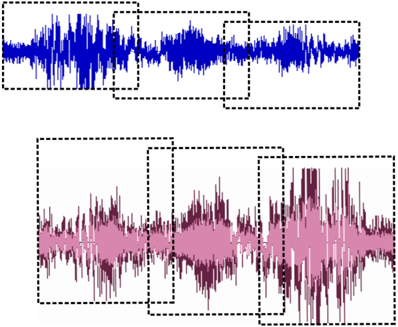
Features and feature extraction

- What is a feature and why do we need it?
  - Captured data might be hard to interpret
  - Many aspects can be contained in a single data stream
  - Example: Audio
    - Loudness
    - Energy on frequency bands
    - Zero crossings
    - Direction changes



Feature subset-selection

- Pre-process data
  - Framing
  - Normalisation



Feature subset-selection

- Pre-process data
  - Framing
  - Normalisation

Domain knowledge?
-> better set of
ad-hoc features

Features commensurate?
-> normalise

Pruning of input required?
-> if no, create disjunctive
features or weithted
sums of features

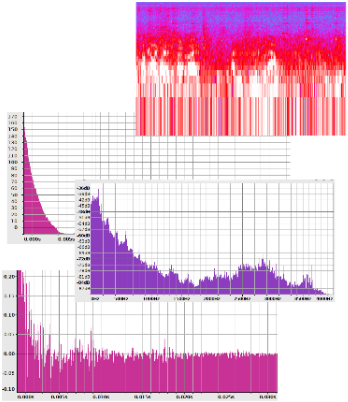
Independent features?
-> construct conjunctive features
or products of features

Is the data noisy?
-> detect outlier examples

Do you know what to do first?
-> If not, use a linear predictor

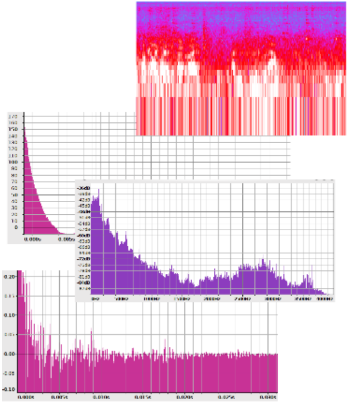
Feature extraction

- Identify meaningful features
  - remove irrelevant/redundant features



Feature extraction

- Identify meaningful features
    - remove irrelevant/redundant features
- Features can be contradictory!



Feature subset-selection

Simple ranking of features with correlation coefficients

Example: Pearson Correlation Coefficient

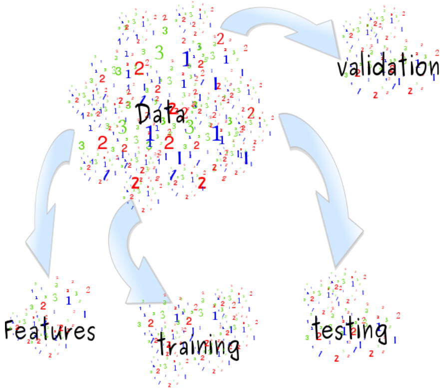
$$\varrho(X, Y) = \frac{\mathrm{Cov}(X, Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}} \tag{1}$$

- Identifies linear relation between input variables x_i and an output y

Feature subset-selection

How to do reasonable feature selection

- Utilise dedicated data sets for feature selection and classifier training
- Pay attention that a single raw-data sample could not impact features in both these sets
- Don't train the features on the training- or test-data-set

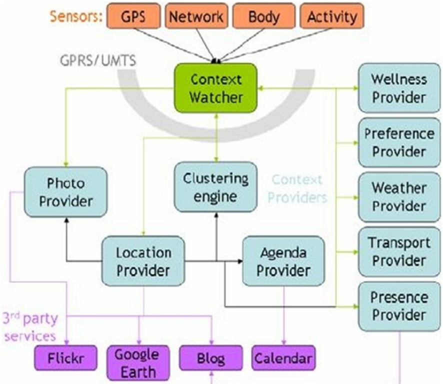


Examples and case studies: Media Cup

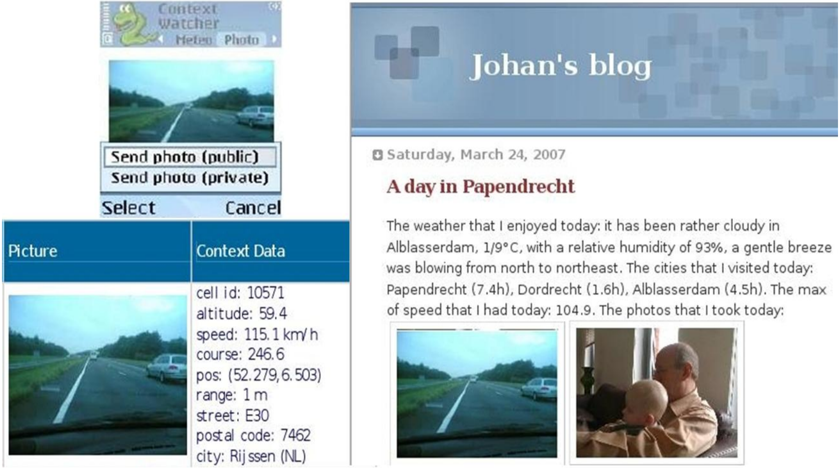
- Media Cup: Context recognition
  - Activity: Trigger sleep mode (save energy)
  - Level of activity
  - Own context: Object movement, person is nervous, specific handling of objects
  - Environmental context: Vibration, earthquake
- Sensor: Ballswitch
  - (nearly) no quiescent current
  - Various types, filled with gas/liquid
  - e.g. Acceleration with fixed value (liquid)
  - Vibration (filled with gas)

Examples and case studies: Context Watcher

- Context Watcher
  - Location
    - GSM cell-ID; GPS
  - Mood
    - user input
  - Activity
    - calender based
  - Bio-data
    - heart and foot sensors
  - Weather
    - location based over internet
  - Photo/picture
    - camera



Examples and case studies: Context Watcher



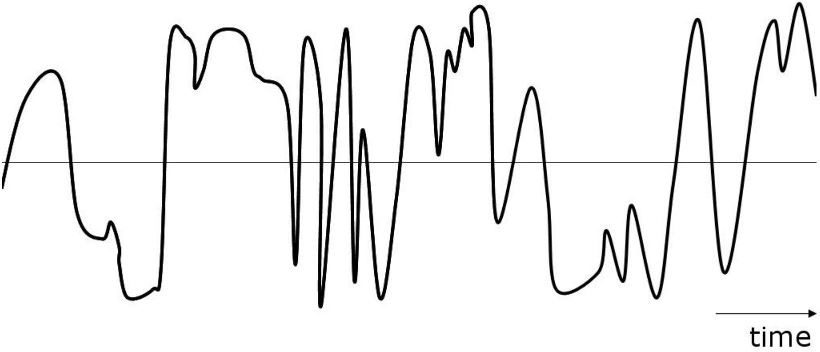
Examples and case studies: TEA

TEA-Audio

- Requirements
  - Restricted memory space
  - Computing power restricted
- Benefit
  - Many sensors $\rightarrow$ Many features
- Example approach
  - Utilise time domain (no transformation)
  - Utilise statistic measures
  - Feature extraction based on small amount of data

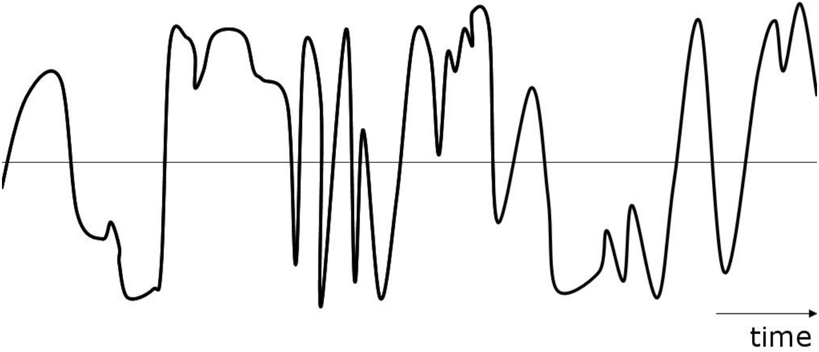
Examples and case studies: TEA

- Audio data in time domain



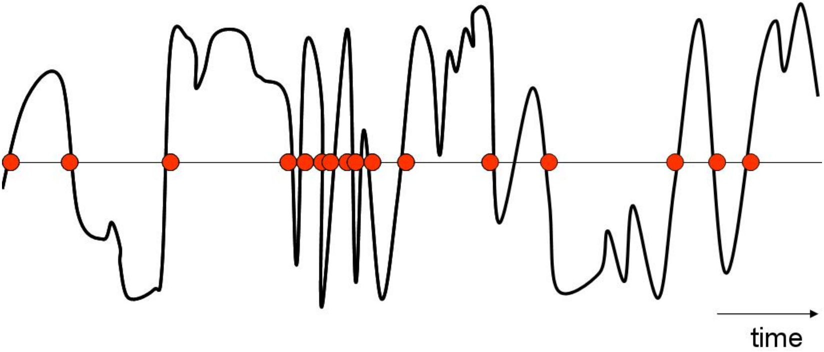
Examples and case studies: TEA

- Audio data in time domain
- Features?



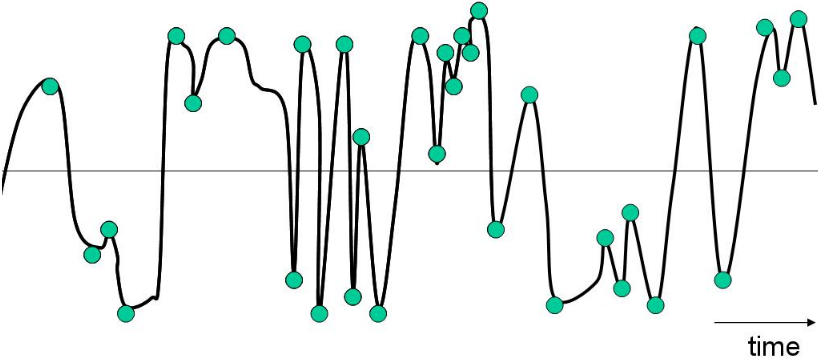
Examples and case studies: TEA

- Count zero crossings
- Distance between zero crossings



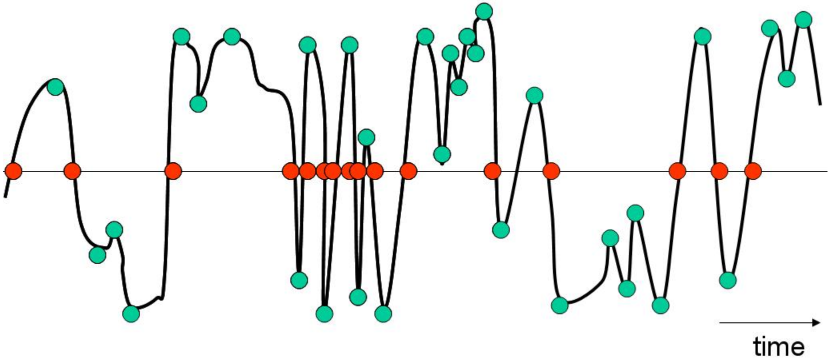
Examples and case studies: TEA

- Count of direction changes



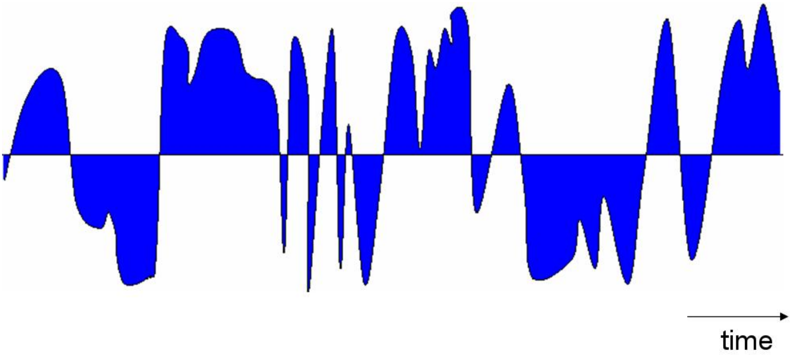
Examples and case studies: TEA

- ratio: $\frac{\text{direction changes}}{\text{zero crossings}}$



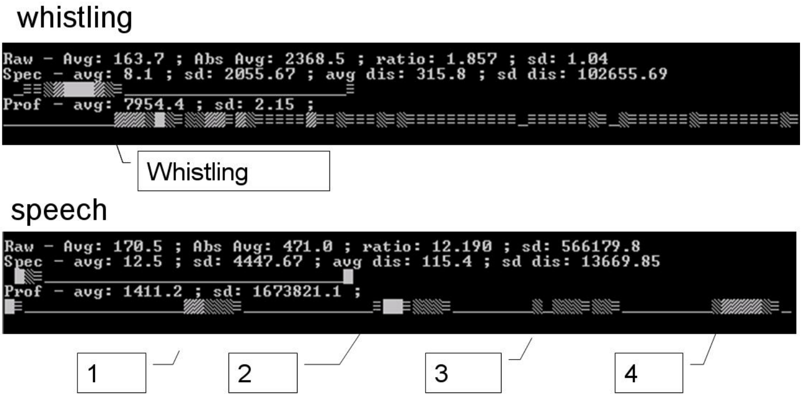
Examples and case studies: TEA

- Integral



Examples and case studies: TEA

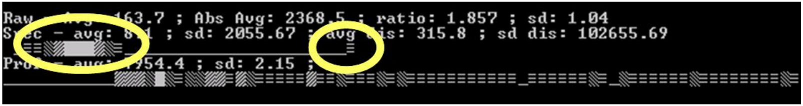
- Several chunks for speech



Examples and case studies: TEA

- Distance between zero crossings: distinct behaviour of oscillation at start and end

whistling

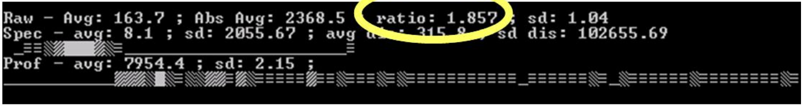


speech

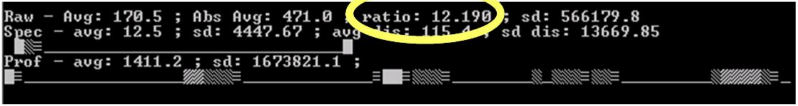
Examples and case studies: TEA

- Distinct ratio: $\frac{\text{zero crossings}}{\text{direction changes}}$

whistling



speech

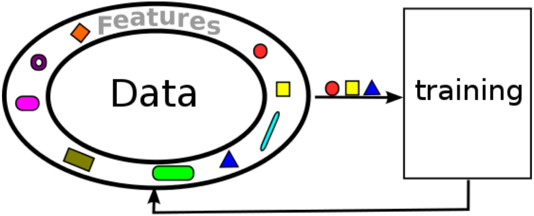


Examples and case studies: TEA

- Significant change in standard deviation of chunks

whistling

speech



Continued next week ...

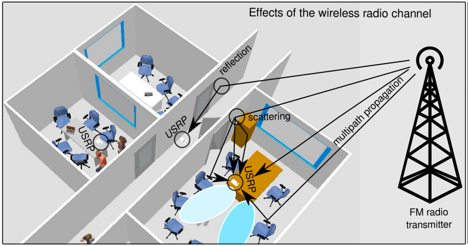
Outline

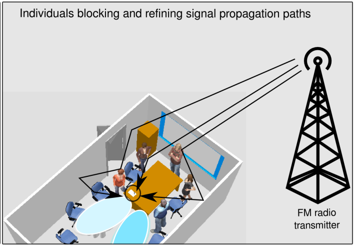
RF-based activity recognition

During propagation, radio signals experience a multitude of effects due to the environment

Can we learn about the environment from the signal evolution observed at a receiver?

- Multi-path propagation
- Scattering
- Reflection
- Blocking of signal paths

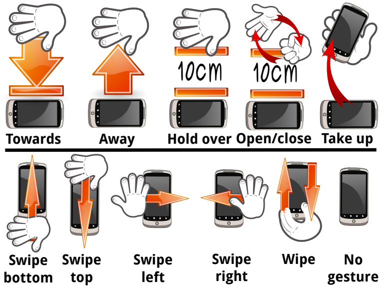




RF-based activity recognition

Sensewaves Video

RF-based activity recognition



Classification

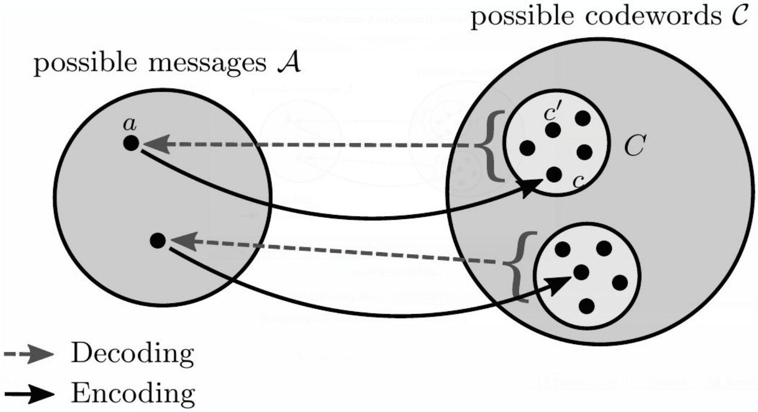
	Aw	Ho	To	No	Op	Sr	St	recall
Aw	**.7**	.02			.06	.09	.13	.70
Ho	.03	**.28**	.22	.05	.2	.16	.06	.28
To		.09	**.76**	.07	.06		.02	.76
No		.05	.06	**.73**	.14	.01	.01	.73
Op	.01	.15	.1	.14	**.49**	.04	.07	.49
Sr	.02	.01		.01	.06	**.83**	.07	.83
St	.12	.03	.01		.05	.14	**.65**	.65
prec	.795	.444	.661	.730	.462	.654	.644	

Classification

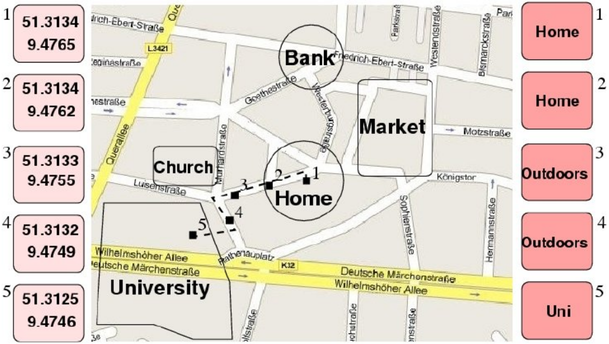
round truth	Away	Towards	No gesture	S. top	recall
Away	**.83**			.17	.83
Towards		**.88**	.09	.03	.88
No gesture	.01	.05	**.92**	.02	.92
S. top	.14	.02	.02	**.82**	.82
precision	.847	.926	.893	.788	

Context-based security

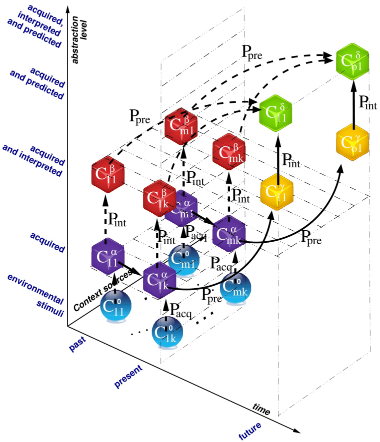
Context-based security

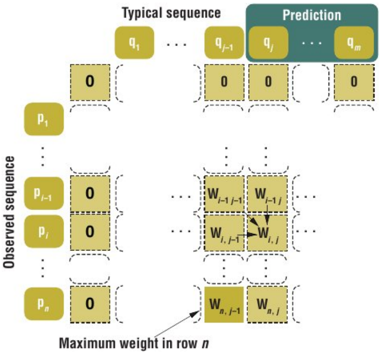


Context prediction

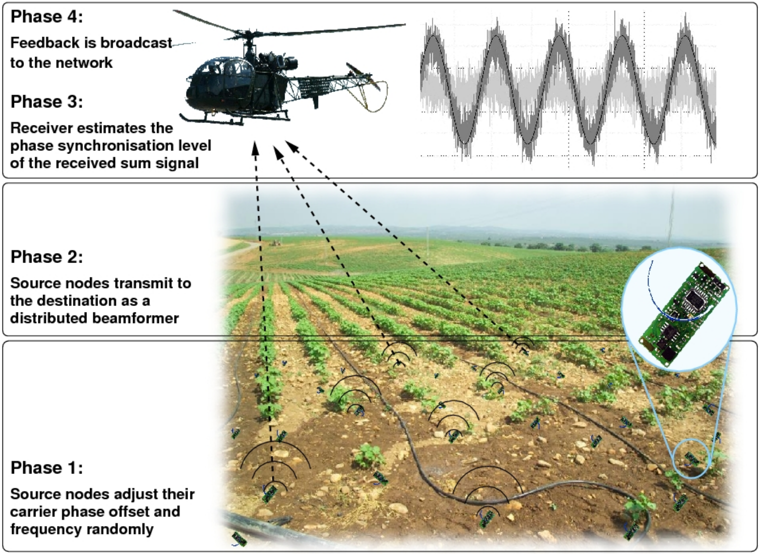


Context prediction





Collaborative transmission



Collaborative transmission / IoT

Outline

Questions?

Stephan Sigg
stephan.sigg@cs.uni-goettingen.de

Literature

- C.M. Bishop: Pattern recognition and machine learning, Springer, 2007.
- R.O. Duda, P.E. Hart, D.G. Stork: Pattern Classification, Wiley, 2001.