Communication Technologies 1 (CT1)

Machine Learning

Introduction

Michel Morold

Lecture in SS 2019
25.04.2019

Contents

ComTec

- Organization
- Introduction
  - Data Mining and Machine Learning
  - Activity Recognition
- Topics for this lecture
- Demonstration

Organization: About us

ComTec

- Prof. Klaus David
  - Dr. Immanuel König
  - M.Sc. Michel Morold
  - M.Sc. Christoph Anderson
  - M.Sc. Judith Heinisch
  - M.Sc. Marek Bachmann
  - ComTec colleagues

- Homepage: https://www.comtec.eecs.uni-kassel.de

About you

ComTec

- You
  - Name
  - Major
  - Previous major
  - Programming skills
  - …

Important: Lecture Registration

ComTec

- A registration for this lecture is **mandatory** and done via HISPOS:
  - https://portal.uni-kassel.de/qisserver/rds?state=verpublish&status=init&vmfile=no&publishid=157040&moduleCall=webInfo&publishConfFile=webInfo&publishSubDir=veranstaltung
- Max. number of students for this lecture: 20
- **First come, first serve**

Organization

ComTec

- Lecture
  - Presentation: 5 topics related to machine learning applications and algorithms
- Exercise (group work) / Lab Training (group work)
  - Scientific paper: 5 pages IEEE double column
  - Programming: task related to algorithms or data processing
- Exam (group work)
  - Presentation
  - Topics related to your paper

Organization: Grade

ComTec

- Group work
  - Paper (and program, depending on specific task) (IEEE, 5 pages)
  - 2-4 Students
  - Presentation
  - Group grade (individual grades possible)
  - **Important:** Create time schedule for your group work and mark individual workload of every group member in your paper (and code, if applicable)

    -> submit time plan to your supervisor until 20.06.2019!)

Organization: Grade

ComTec

- Individual Bonus
  - Submit lecture summary via e-mail (**max. 2 slides**) to **michel.morold@comtec.eecs.uni-kassel.de** **until 11:59 p.m. of the day after the lecture (= Friday, 11:59 p.m.)**
  - Use PowerPoint template in Moodle
  - Presentation (3-5 minutes) if selected at the beginning of the next lecture
  - Use file name “summary_<your last name>_<title>_<yyyymmdd>.ppt(x)”
  - Minimum 4 reasonable summaries = improve grade by +0.3/0.4

Start on 02.05.2019

Grade: Evaluation criteria

Task	Criterion	Achieved	Not achieved
Paper (5 pages IEEE double column, use template in Moodle)	Technically comprehensive, clear structure	+/-	+/-
	Complete, content free from errors	+/-	+/-
	Shape and impression	+/-	+/-
	Grammar, expression, spelling	+/-	+/-
	Cite 3 (IEEE / ACM) publications related to own paper		-0.3/0.4
	Autonomy of the group	+/-	+/-
	PLAGIARISM		failed
Presentation (use template in Moodle)	Content, clear structure	+/-	+/-
	Rhetoric, language	+/-	+/-
	Talk time - 15-20 minutes per group (depending on group size)		-0.3/0.4
	Answering questions	+/-	+/-
	No presentation given		failed
Summaries (bonus)	Hand in at least **4 reasonable** summaries and present summary if selected at the beginning of the next lecture	+0.3/0.4	

Grade: Evaluation criteria

ComTec

Team Work

Plagiarism = failed

Organization: Time plan

ComTec

Time	Lecturer	Topic
25.04.2019	MM	Introduction
02.05.2019	JH	Segmentation / Features
09.05.2019	MM	Bayesian Classification
16.05.2019	CA	Decision Trees
23.05.2019	IK	Hidden Markov Models
06.06.2019	MB	Clustering / KNN and presentation of student tasks
06.06.2019 – 20.09.2019		Lab/Programming
20.06.2019		**Deadline submission time/work plan (for supervisor)**
01.08.2019		**Deadline of the draft paper and program**
19.09.2019	All	• **Presentations (10:00 – t.b.a.)** • **Deadline of presentation submission**
23.09.2019		**Deadline of the final paper and program**

Problems

ComTec

- Registration
  - Alexander Bolz: alexander.bolz@comtec.eecs.uni-kassel.de
  - Michel Morold: michel.morold@comtec.eecs.uni-kassel.de
- Topic / task
  - Supervisor (e-mail)
- Appointment
  - Topic/task (mandatory)
  - Feedback of the paper task, preparation of the presentation task (mandatory), but maximum 4

Registration

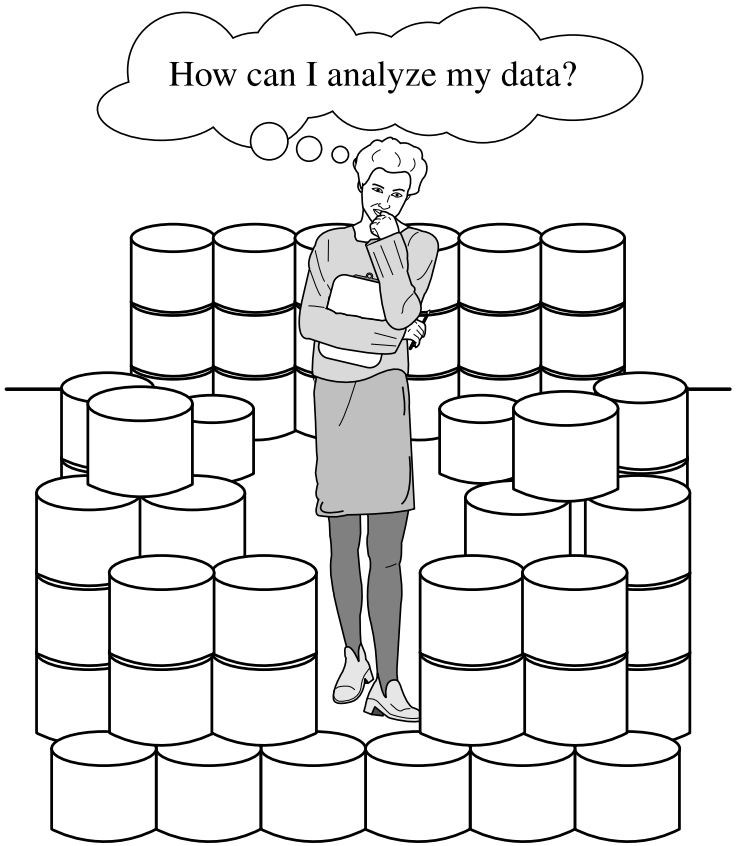
ComTec

- Exam Registration for Lecture
  - Master Informatik (HIS-POS) (FB16-5354)
    - Communication Technologies 1 (PNr. 108010)
  - ECE (OKA):
    - Communication Technology I (5354.1)
    - Laborpraktikum Optische Nachrichtentechnik I (5322.1)
  - **Deadline for lecture registration: TBA**
- Moodle
  - Web: https://moodle.uni-kassel.de/moodle/course/view.php?id=2185
  - Password: **MM-19-CT1**

Data Mining

ComTec

"...due to the wide availability of huge amounts of data and there is an imminent need for turning such data into useful information and knowledge..." [1]



[1]

Data Mining

ComTec

- Data Mining
  - To solve problems by analyzing data already present in database [2]

- Data / Database
  - Google
  - Facebook
  - LinkedIn
  - Apple
  - Amazon
  - Weather Service
  - National Institutes of Health / Hospitals

Data Mining

ComTec

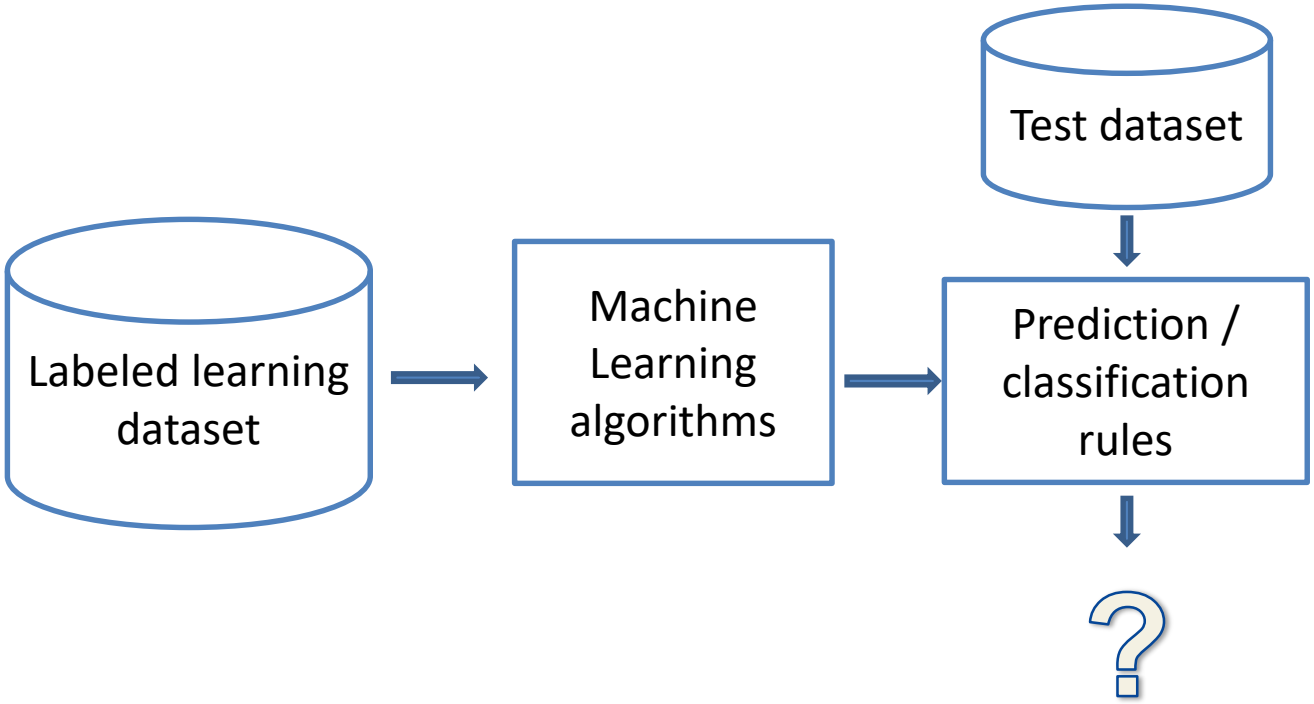
- Data Mining
  - To solve problems by analyzing data already present in database [2]

- Data / Database
  - Google -> how search engines rank
  - Facebook -> who you know
  - LinkedIn -> which job you prefer
  - Apple -> where you are
  - Amazon ->which product you would buy
  - Weather Service -> weather forecast
  - National Institutes of Health / Hospitals -> disease diagnosis

Machine Learning

ComTec

- Machine Learning
  - To develop algorithms for making decisions/predictions from data [3]
  - “Learning is any process through which a system acquires synthetic a posteriori knowledge” [4]



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…

Activity Recognition (AR)

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

- Vision-based AR
- Sensor-based AR

Daily Life Activities

Activity Recognition (AR)

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

- Vision-based AR
- Sensor-based AR

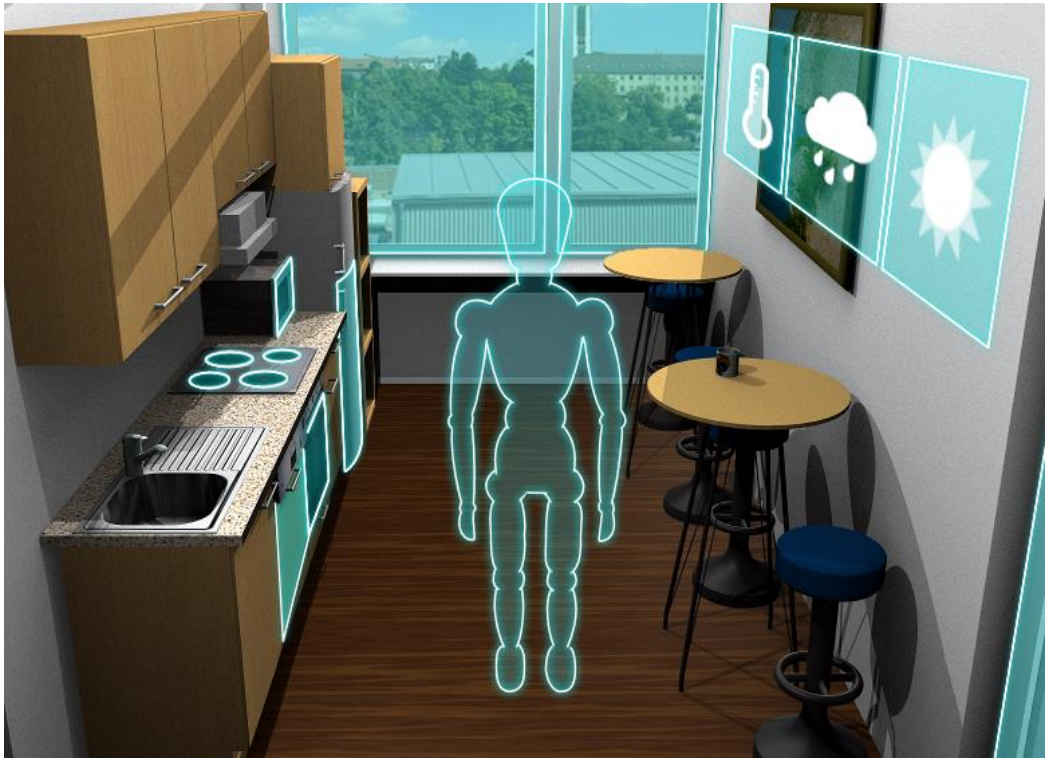
Daily Life Activities

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 object within an environment



ComTec

Sensor-based Activity Recognition

ComTec

Wearable Sensors

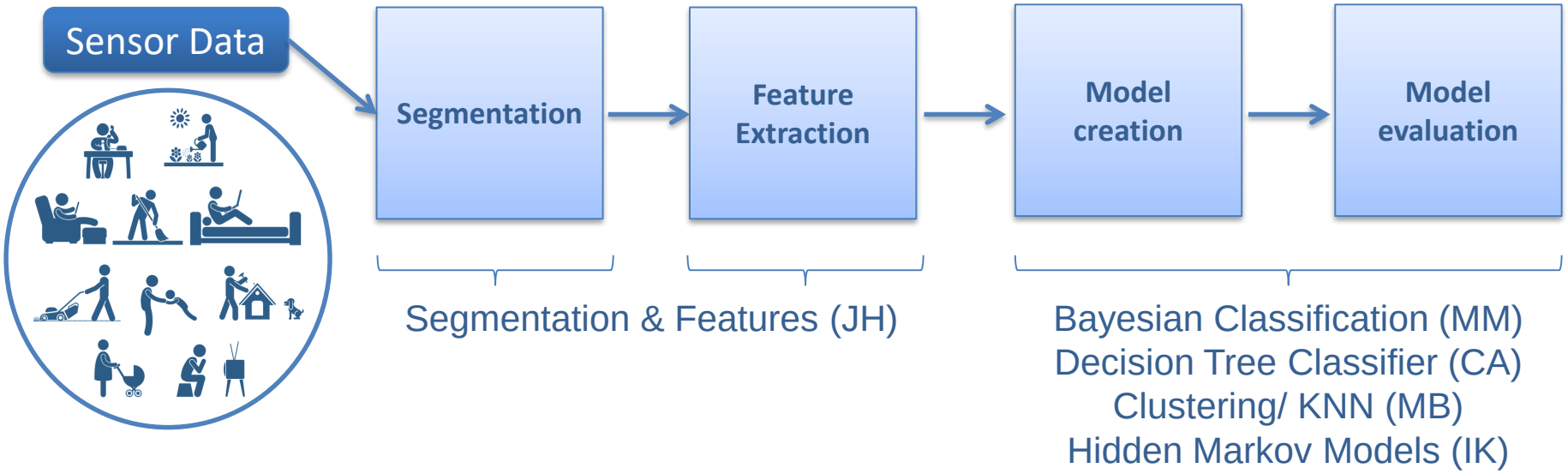
- Can be positioned directly or indirectly on the body
- Sensors be embedded into clothes, mobile devices, etc.
- Position, pulse, skin temperatures
- Mainly physical activities like walking, running or standing

[5]

Topics for this lecture

ComTec

Machine Learning Algorithms for Context Recognition



Demonstration

ComTec

- Android-based Activity Demonstrator
  - Sensors: accelerometer and gyroscope
  - Fixed window size: 1 s
  - Features: mean, variance, minimum, maximum
  - Classification algorithms: J48 (C4.5), Jrip, NaiveBayes

Demonstration: J48 example

ComTec

```
accelerometer-y_variance <= 0.863208
|   accelerometer-z_minimum <= -7.244843: sitting (90.0)
|   accelerometer-z_minimum > -7.244843
|   |   accelerometer-z_minimum <= 1.515518
|   |   |   accelerometer-x_variance <= 0.349223: standing (267.0)
|   |   |   accelerometer-x_variance > 0.349223
|   |   |   |   gyroscope-z_mean <= -0.131582: walking (2.0)
|   |   |   |   gyroscope-z_mean > -0.131582: standing (8.0/1.0)
|   |   accelerometer-z_minimum > 1.515518: sitting (64.0)
accelerometer-y_variance > 0.863208
|   accelerometer-y_variance <= 99.880016
|   |   gyroscope-y_variance <= 0.117366
|   |   |   accelerometer-z_maximum <= -1.103714: sitting (3.0)
|   |   |   accelerometer-z_maximum > -1.103714
|   |   |   |   accelerometer-x_minimum <= -2.573761: walking (5.0/1.0)
|   |   |   |   accelerometer-x_minimum > -2.573761
|   |   |   |   |   accelerometer-y_variance <= 31.575412: standing (6.0)
|   |   |   |   |   accelerometer-y_variance > 31.575412: walking (2.0)
|   |   gyroscope-y_variance > 0.117366
|   |   |   gyroscope-y_minimum <= -4.492218
|   |   |   |   accelerometer-y_variance <= 43.384506: walking (6.0)
|   |   |   |   accelerometer-y_variance > 43.384506: running (3.0)
|   |   |   gyroscope-y_minimum > -4.492218: walking (305.0/5.0)
|   accelerometer-y_variance > 99.880016: running (110.0)
```

References

[1] J. Han and M. Kamber, Data mining: concepts and techniques, 2nd ed. Amsterdam ; Boston : San Francisco, CA: Elsevier ; Morgan Kaufmann, 2006.

[2] I. H. Witten and E. Frank, *Data Mining: Practical Machine Learning Tools and Techniques*, 2nd ed. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2005.

[3] R. Bekkerman, M. Bilenko, and J. Langford *Scaling up Machine Learning: Parallel and Distributed Approaches*, New York, NY, USA, Cambridge University Press, 2011

[4] P. D. Scott, “Learning: the construction of a posteriori knowledge structures,” in Proceedings of the Third AAAI Conference on Artificial Intelligence, 1983, pp. 359–363.

[5] L. Bao and S. S. Intille, “Activity Recognition from User-Annotated Acceleration Data,” in Pervasive Computing: Second International Conference, PERVASIVE 2004, Linz/Vienna, Austria, April 21-23, 2004, pp. 1–17.