Authentication

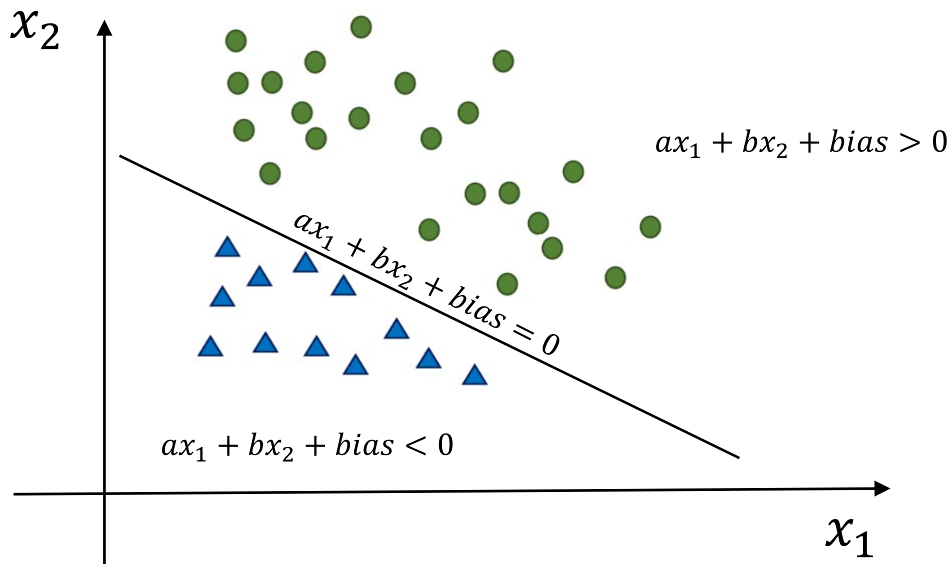
Prof. Dr. Helene Dörksen

helene.doerksen@th-owl.de

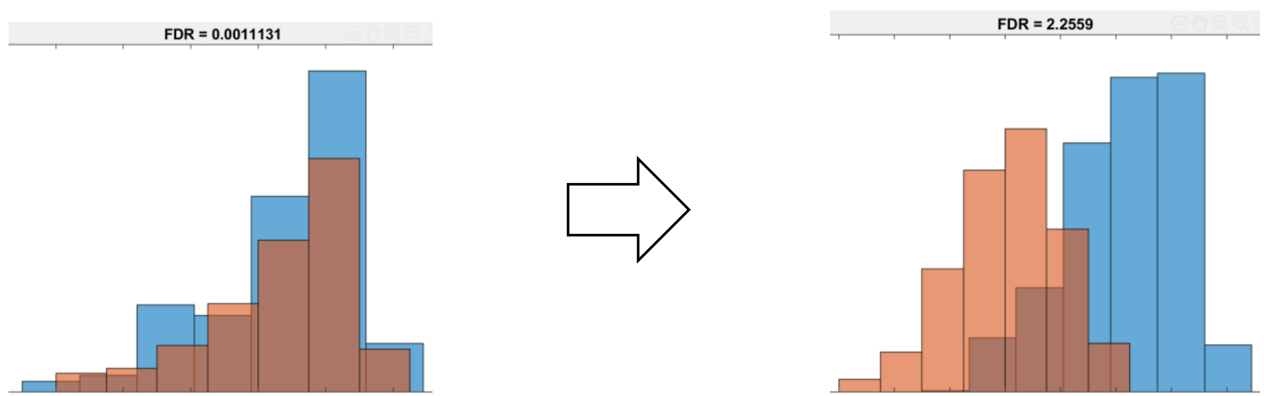
We need to know about

Linear Classifiers, e.g.

- support-vector-machine
- LDA



Principles of feature extraction = combining the existing feature set into a smaller set of new, more informative features



Lecture 9:

Combinatorial Refinement of Feature Weighting for Linear Classification

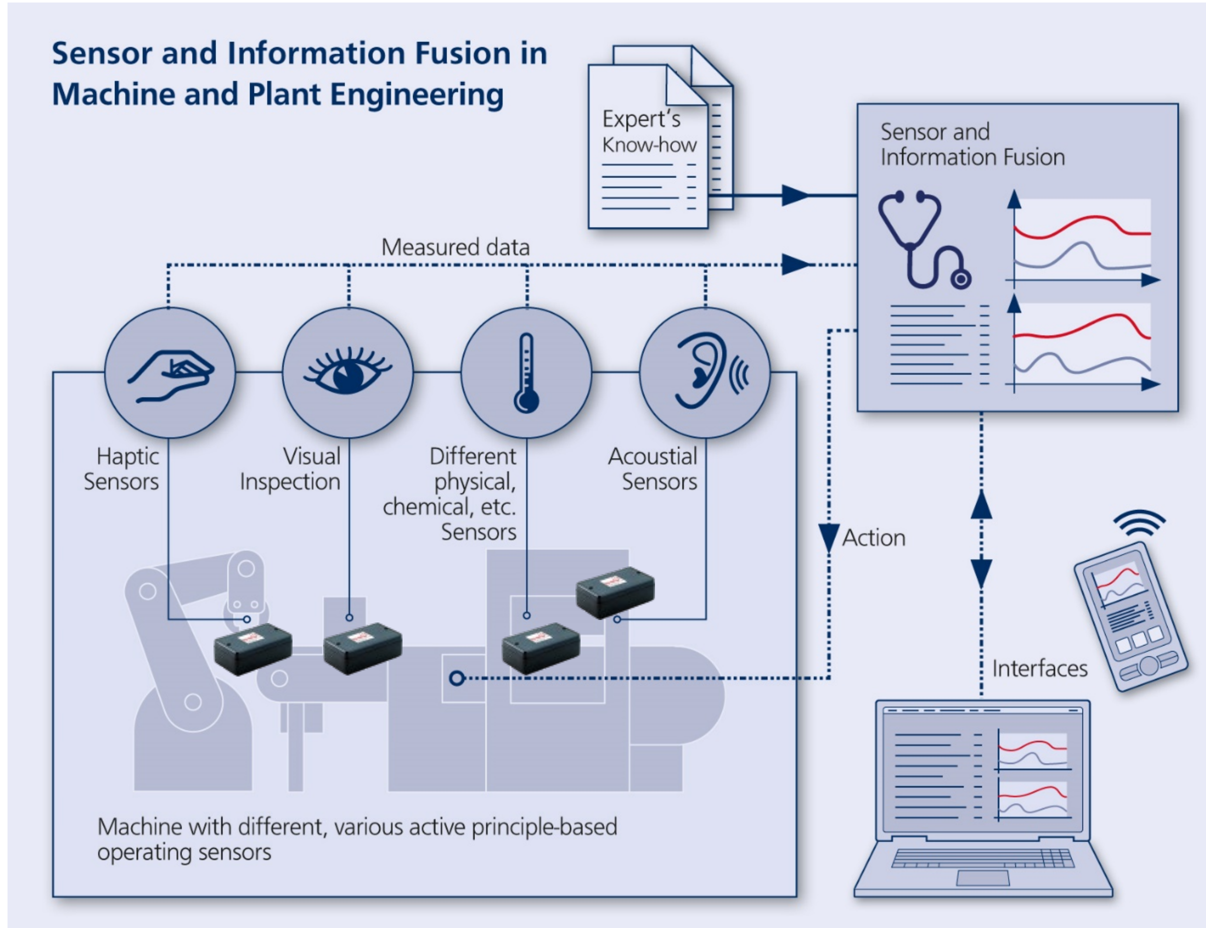
19th IEEE International Conference on
Emerging Technologies and Factory Automation, Barcelona, Spain, 16.-19. September 2014

Outline

- Introduction
- Related Work:
  - Dimensionality Reduction
  - Multiple Learners
- Combinatorial Refinement *(**ComRef**)* Approach for Classification
- Generalization Ability of ***ComRef***
- Conclusion and Outlook

Introduction

Automized Container Terminals
in Harbour of Hamburg
(Germany)

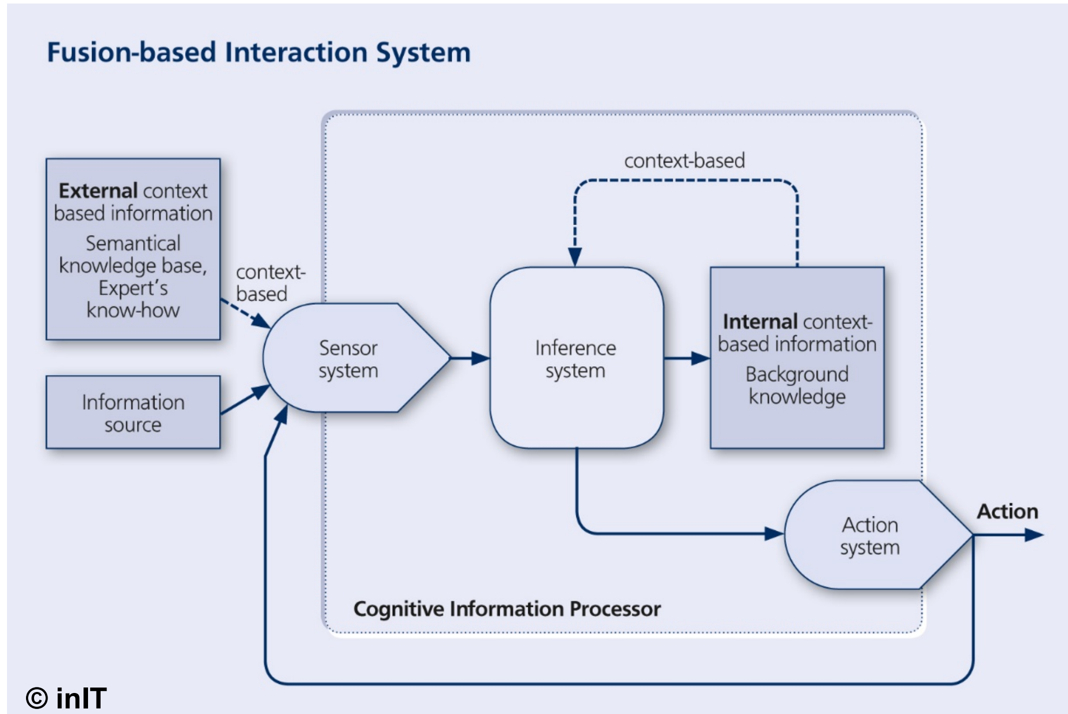


Introduction

A Cyber-physical system (**CPS**) is a system of collaborating elements

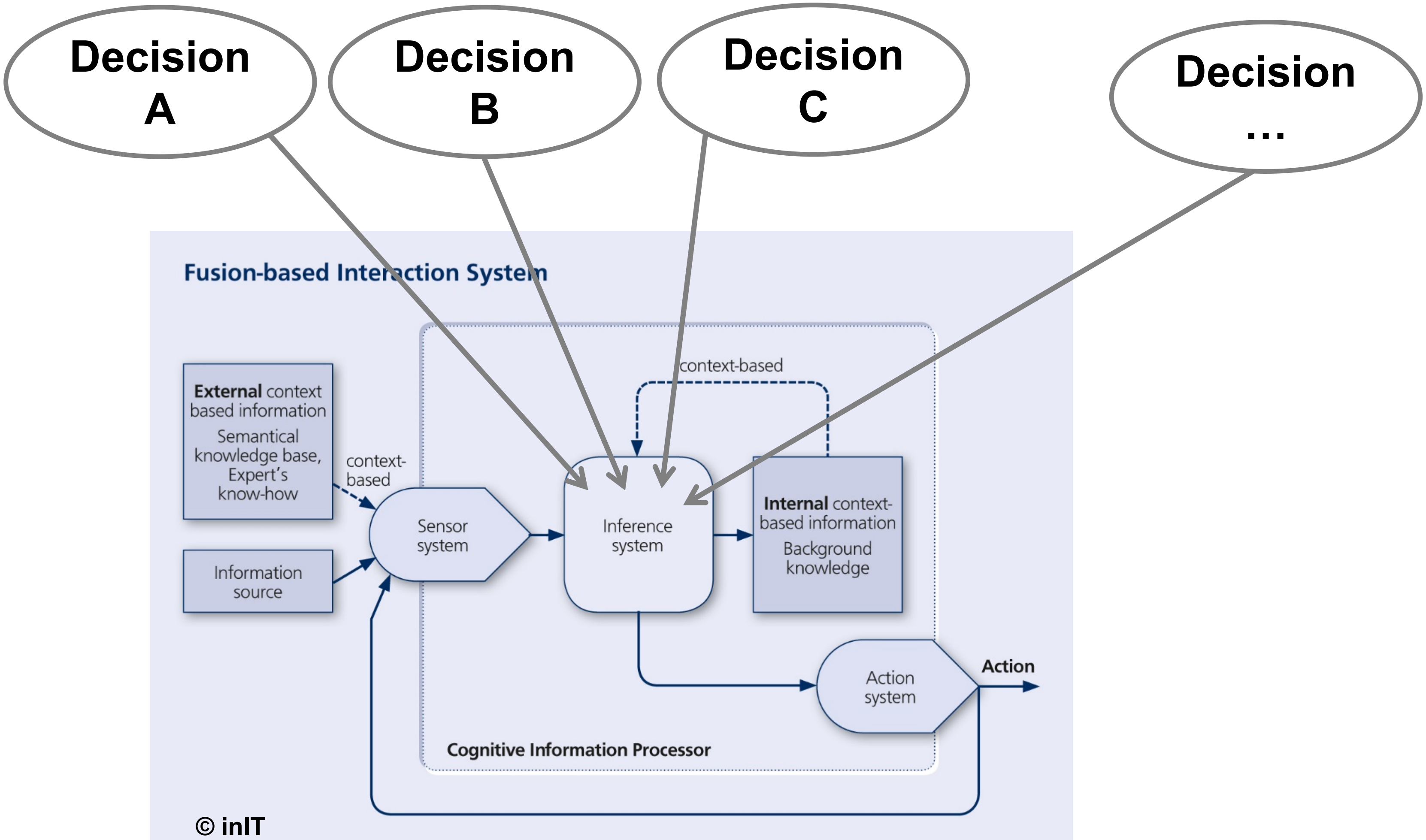
Applications:

- Automotive
- Manufacturing
- Transportation
- Healthcare
- …

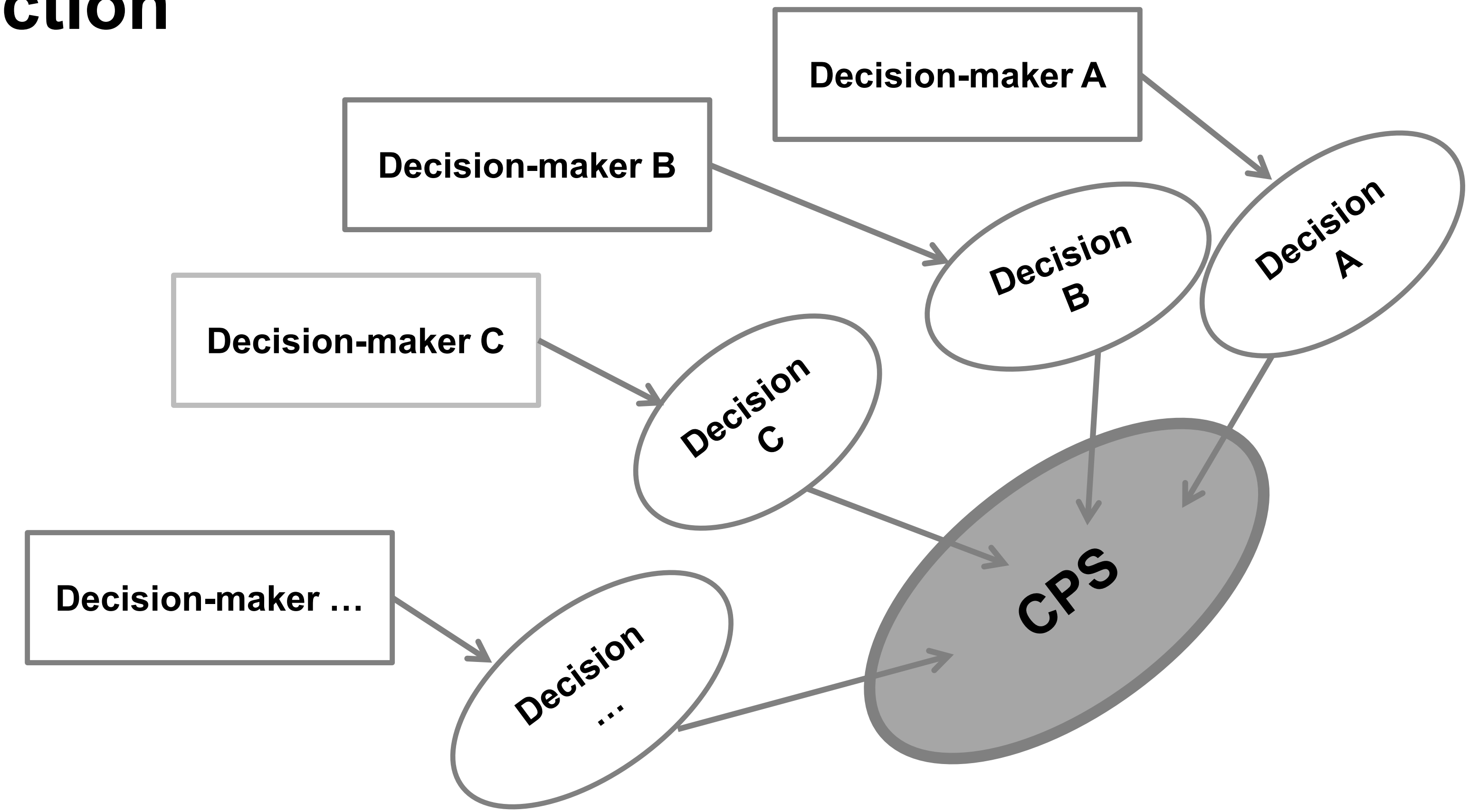


CPS work context-specific, adaptive, partly autonomous, automized, multi-functional and multi-sensory
→ **high degree of complexity**

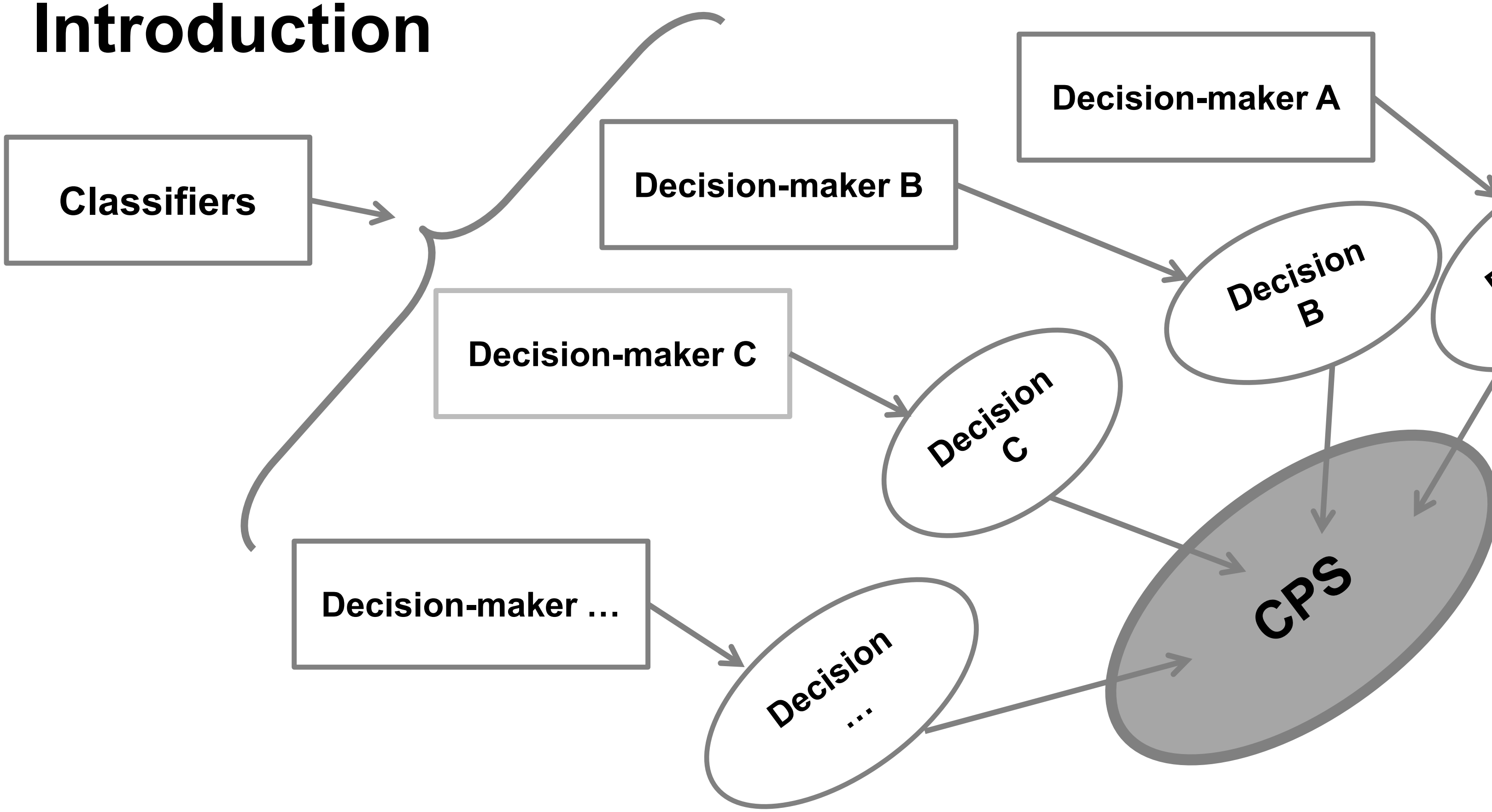
Introduction



Introduction



Introduction



Basis for **good generalisation** of a classifier [1]:

- Simple models (*Occam's razor* principle)
- Regularisation (e.g. large classification margin like SVM)

[1] I. Guyon, S. Gunn, M. Nikravesh, and L. A. Zadeh, Feature Extraction: Foundations and Applications (Studies in Fuzziness and Soft Computing). Secaucus and NJ and USA: Springer-Verlag New York, Inc, 2006.

Introduction

<u>Primary</u> basics for **good generalisation** [1)] of a classifier:

- Simple models (*Occam‘s razor* principle)
- Regularisation (e.g. large classification margin like SVM)

<u>Secondary</u> basics (methods) might be **good for generalisation** [1)] of a classifier:

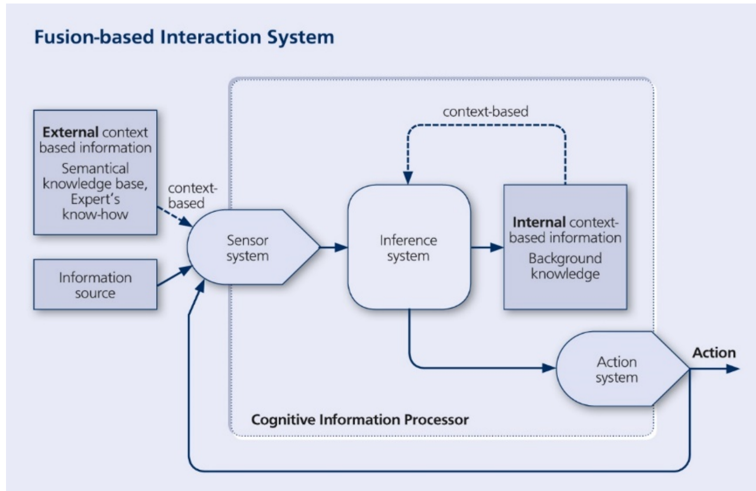
- Dimensionality Reduction
- Multiple Classifiers
- …

[1)] E. Alpaydın, Introduction to Machine Learning, 2nd ed. Cambridge: The MIT Press, 2010.

Introduction Summary

Automized Container Terminals
in Harbour of Hamburg
(Germany)

Implementation System



Actions are based on classifiers

Basics for **good generalisation of classifier:**

- Simple models (*Occam's razor* principle)
- Regularisation (e.g. large classification margin like SVM)

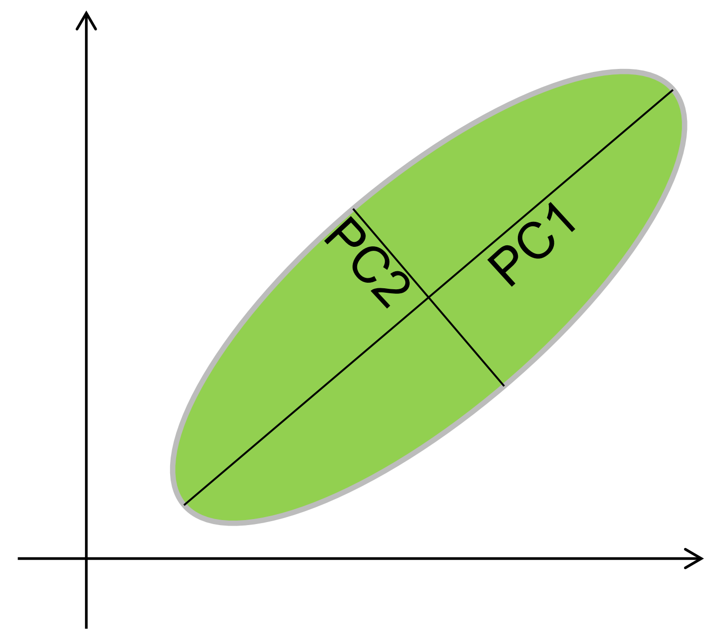
Methods:

- Dimensionality Reduction
- Multiple Classifiers

Related Work: Dimensionality Reduction [1)]

DR (or **feature selection**) = Selecting subsets of most useful features and ignoring the rest

- Principal Component Analysis (PCA)
- Linear Discriminant Analysis (LDA)
- Factor Analysis (FA)
- …

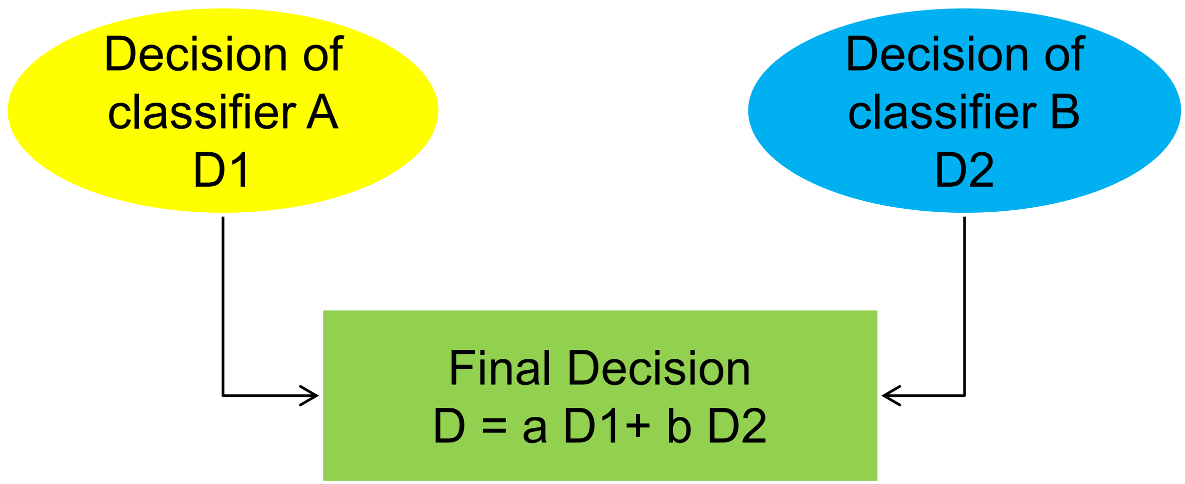


1) E. Alpaydın, Introduction to Machine Learning, 2nd ed. Cambridge: The MIT Press, 2010.

Related Work: Multiple Classifiers [1]

- Bag-of-Classifiers
- Bag-of-Features
- Ensemble Learners
- …

Classifiers complement each other to obtain higher accuracy of decision



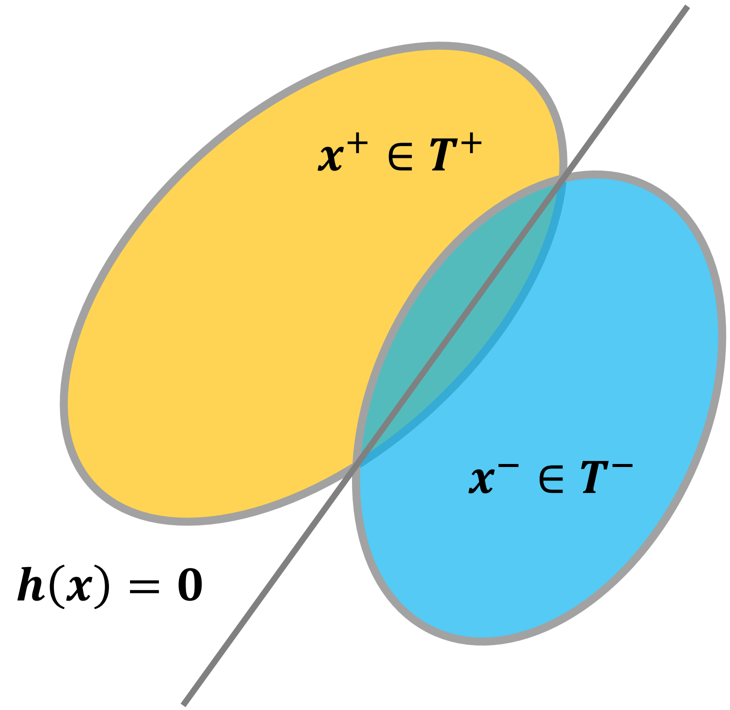
[1] L.I. Kuncheva, Combining Pattern Classifiers: Methods and Algorithms. Wiley-Interscience, 2004.

ComRef Approach for Classification

Workflow for classifier training:

- Consider suitable classification techniques
- Perform generalisation test(s) to evaluate the accuracies (and their standard deviations) of classification rule
- Choose the technique with highest generalisation ability as final classifier

Measurements (Features) in $\mathbb{R}^d$



Classification Rule:

$$x \in T^+ \quad if \quad h(x) \geq 0$$
$$x \in T^- \quad if \quad h(x) < 0$$

ComRef Approach for Classification

All features are equally relevant for the classification boundary calculation:

$$\boldsymbol{h(x) = a_1x_1 + a_2x_2 + \cdots + a_i\,x_i + a_{i+1}x_{i+1} + \cdots + a_dx_d + c_h}$$

Principles of the *feature extraction* [1)] :

- a helpful feature may be **less relevant** by itself;
- combination of individually irrelevant features **may become relevant**

1) I. Guyon, S. Gunn, M. Nikravesh, and L. A. Zadeh, Feature Extraction: Foundations and Applications (Studies in Fuzziness and Soft Computing). Secaucus and NJ and USA: Springer-Verlag New York, Inc, 2006.

ComRef Approach for Classification

$$h(x) = a_1 x_1 + a_2 x_2 + \cdots + a_i\, x_i + a_{i+1} x_{i+1} + \cdots + a_d x_d + c_h$$

summands: $a_j x_j$

fusion of summands: $\sum_{j \in I} a_j x_j$

<u>Idea:</u>

- a helpful summand may be **less relevant** by itself
- a **fusion** of individually irrelevant **summands** **may become relevant**

ComRef Approach for Classification

$$h(x) = \boxed{a_1 x_1 + a_2 x_2 + \cdots + a_i\, x_i} + \boxed{a_{i+1} x_{i+1} + \cdots + a_d x_d} + c_h \qquad (I)$$

fusion of summands $\underbrace{1, \cdots, i}$ fusion of summands $\underbrace{i+1, \cdots, d}$

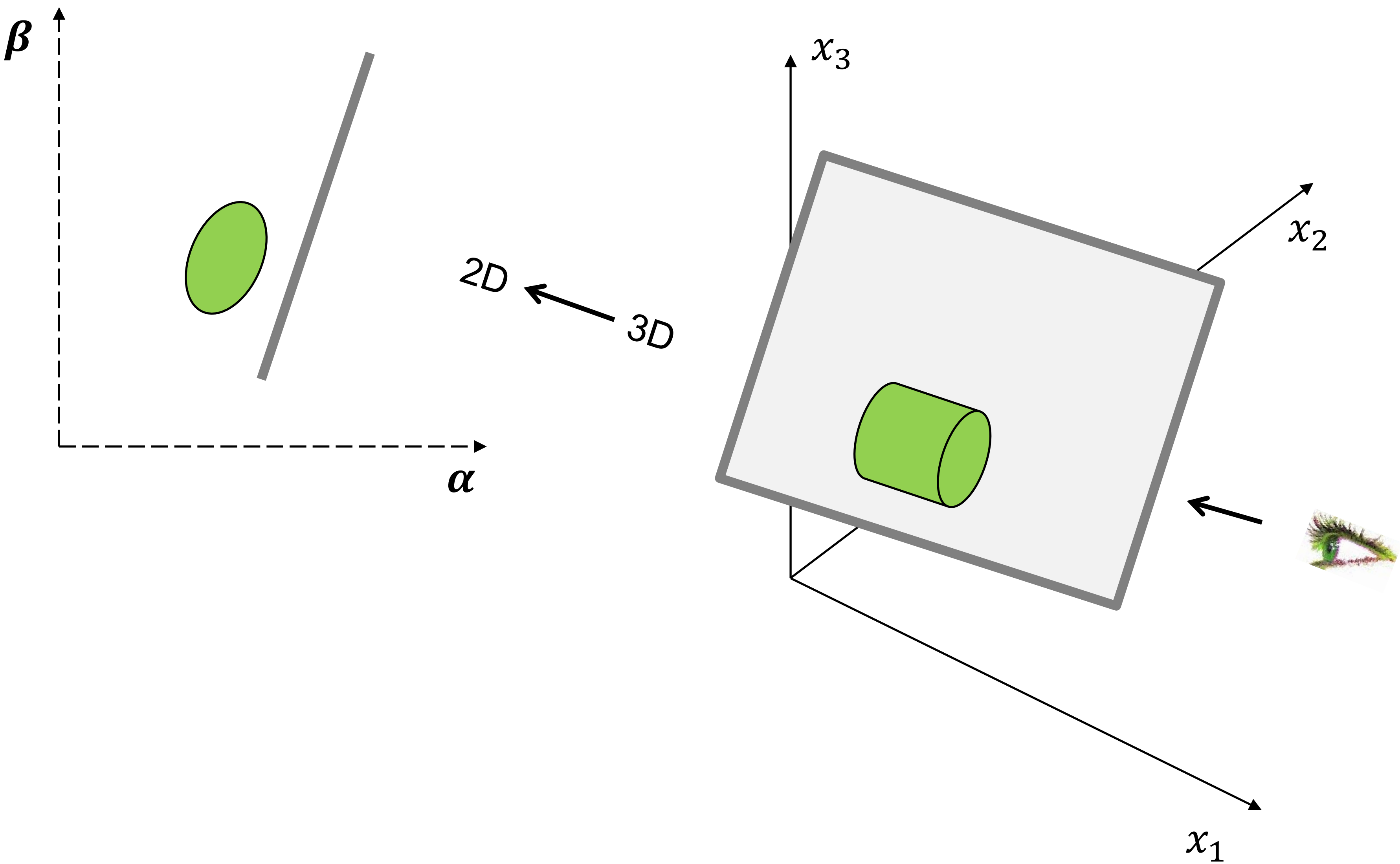
$$\underbrace{\sum_{j=1}^{i} a_j x_j} \qquad \underbrace{\sum_{j=i+1}^{d} a_j x_j}$$

$$\underbrace{\sum_{j \in I_1} a_j x_j \qquad \sum_{j \in I_2} a_j x_j}$$

$$h(x) = \underbrace{\boxed{\sum_{j \in I_1} a_j\, x_j}}_{\alpha} + \underbrace{\boxed{\sum_{j \in I_2} a_j x_j}}_{\beta} + c_h \qquad (II)$$

Dimensionality Reduction: d summands $\rightarrow$ 2 summands,
but classification rules of (I) and (II) are equivalent!

ComRef Approach for Classification

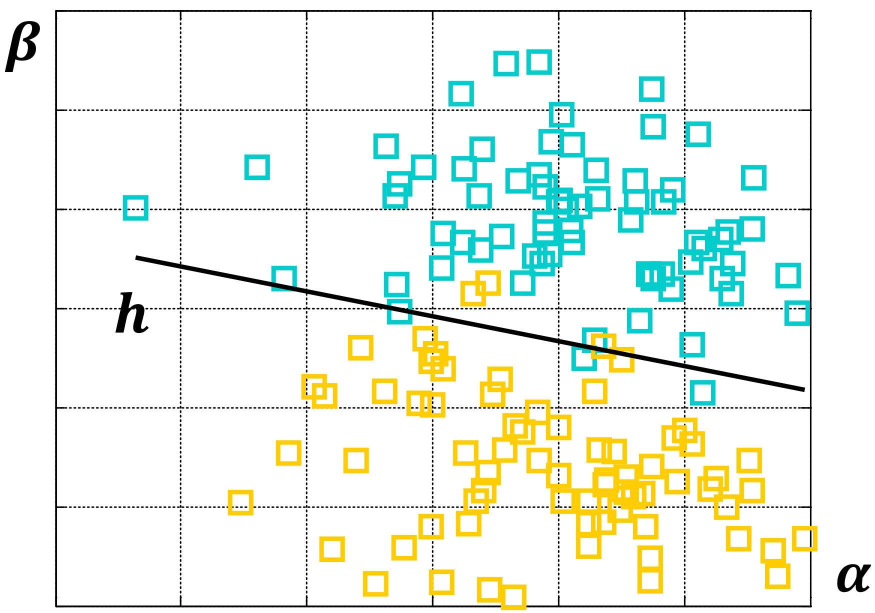


ComRef Approach for Classification

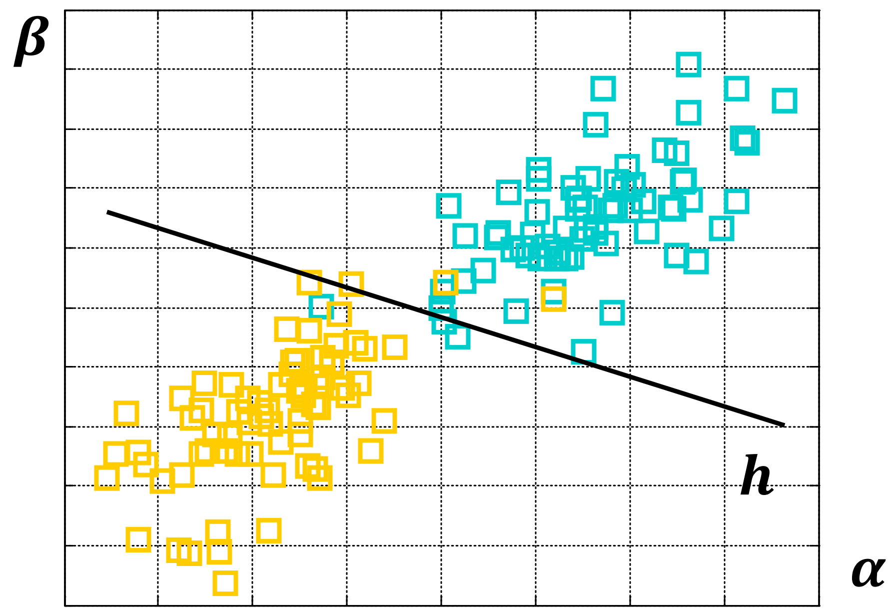
SVM for classes □ and □ of ***SEEDS*** [1)]:

$$h(x) = \sum_{i=1}^{d} a_i x_i + c_h \quad (d = 7)$$

Dimensionality Reduction by $I_1 = \{1, 2, 3, 7\}$ and $I_2 = \{4, 5, 6\}$



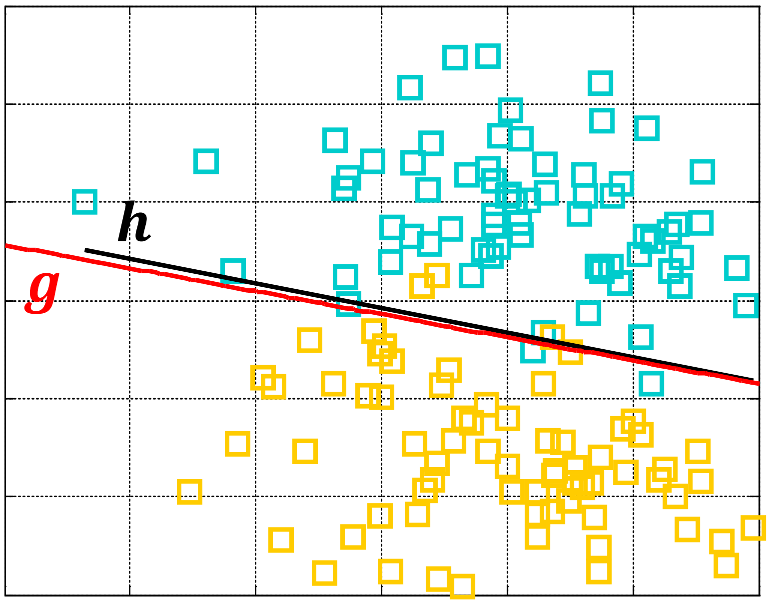
Dimensionality Reduction by $I_1 = \{4, 7\}$ and $I_2 = \{1, 2, 3, 5, 6\}$



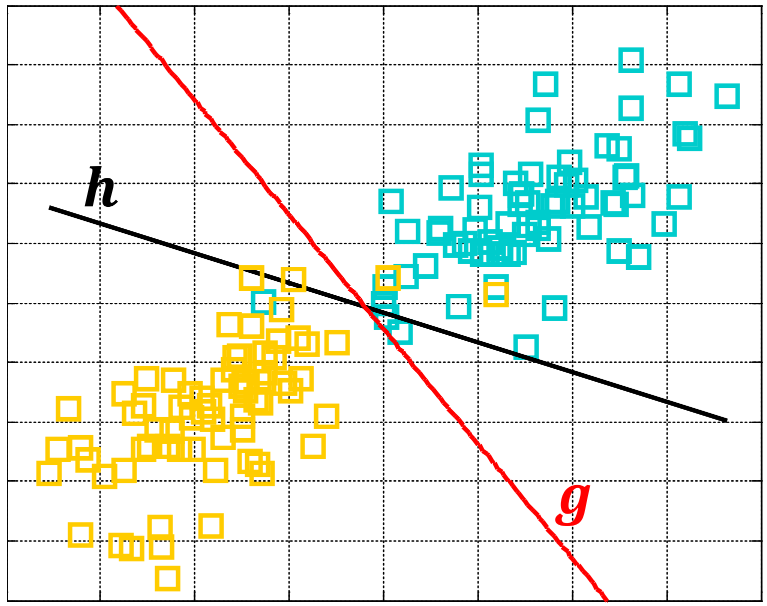
classification rate of h for both reductions: 95.71%

[1)] UCI Machine Learning Repository: https://archive.ics.uci.edu/ml/index.php

ComRef Approach for Classification



classification rate of g: 96.43%



classification rate of g: 97.85%

Combination of classifiers:
Multiple Classifier: initial (linear) classifier h + refinement (linear) classifier g)

We are able to improve accuracy without increasing complexity!

ComRef Approach for Classification

Short summary:

- By fusions of summands **Dimensionality Reductions** are possible
- Combinations of classifiers: **Multiple Classifier** = initial classifier + refinement classifier, are possible
- **Improvement of accuracy** without increasing complexity is possible

ComRef Approach for Classification

***ComRef* Workflow Example:**

1. Choose initial classifier $\boldsymbol{h}$
2. Reduce the dimensionality by fusions of summands of $\boldsymbol{h}$
3. Choose refinement classifier/classifiers $\boldsymbol{g}$ for reduced spaces
4. Select fusion of summands with highest performance resp. $\boldsymbol{g}$

ComRef Approach for Classification

Fusions of summands lead to **dimensionality reductions** down to spaces:

$\boldsymbol{d-1} \leftarrow$ $\boldsymbol{I_1}$ consists of **2** summands, $\boldsymbol{I_2}$ is rest; complexity of calculation $\boldsymbol{O(d^2)}$ [1)]

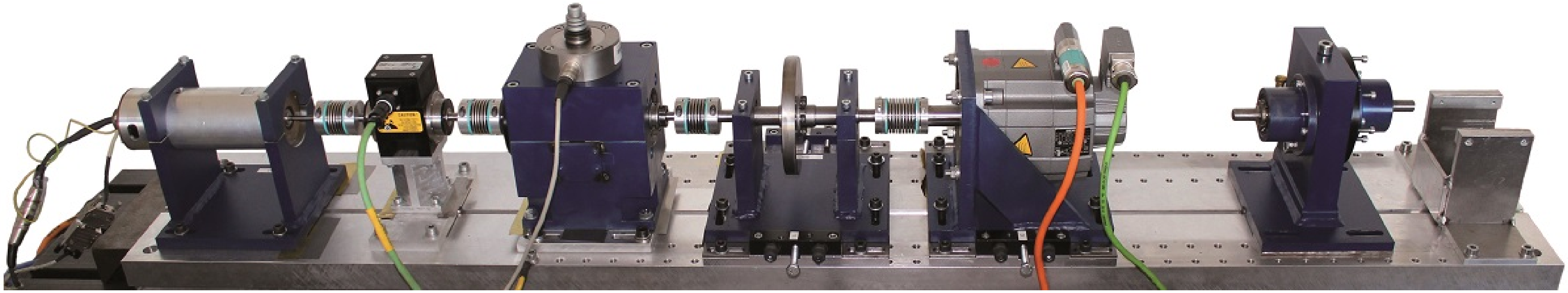
$$\left.\begin{array}{c} \boldsymbol{d-2} \\ \boldsymbol{d-3} \\ \vdots \\ \boldsymbol{3} \end{array}\right\}$$ **combinatorial task**

$\boldsymbol{2} \leftarrow$ $\boldsymbol{I_1}$ is complementary to $\boldsymbol{I_2}$; complexity of calculation $\boldsymbol{O(2^d)}$ [1)]

[1)] H. Dörksen and V. Lohweg, “Combinatorial Refinement of Feature Weighting for Linear Classification,” in 19th IEEE Int. Conf. on Emerging Technologies and Factory Automation (ETFA 2014), 2014.

Generalisation Ability of *Comref*

Motor Drive Diagnosis: motors as CPS which are networked



e.g. one motor with **53190** measurement-attributes for **72** features

- **72** features lead to 100% accuracy by linear SVM

but:

- **9** features (of **72**) leading to $\sim 100\%$ (99.9933%) accuracy by linear SVM (without ***ComRef***) are found
- **5** features (of **72**) leading to 100% accuracy by linear SVM (with ***ComRef***) are found

Conclusion and Outlook

- ✓ In **reduced space** improvement of accuracy is possible without increasing complexity

- ✓ Combination of classifiers (**multiple classifier**: initial classifier – refinement classifier) is possible

- ✓ **Many examples** are found for which ***ComRef*** **improves** the **accuracies**

- ✓ ***ComRef*** **did not dis-improve** accuracies of **no one** of tested examples

→ **Combinatorial task** for spaces of dimensions $\boldsymbol{d-2, \ldots, 3}$ has to be solved as efficient as possible

→ Problem of **selection of fusion of summands** with highest refinement classification accuracy has to be solved efficiently

Summary

Combinatorial Refinement ***(ComRef)*** Approach for Classification

→ is able to improve accuracy without increasing complexity!

→ is based on principles of:

- Dimensionality Reduction
- Multiple Learners

improvement →

Automized Container Terminals in Harbour of Hamburg (Germany)

Homework: Exercises and Labs

for the next week prepare practical exercises and labs from
Exercises Lec 9 (you will find it in the download area)

Online-Exercises coming now

Exercises Lec 8 will be discussed