Eberhard Karls Universität Tübingen
Wilhelm-Schickard-Institut für Informatik
Lehrstuhl Technische Informatik

# Feature Selection for High-Dimensional Industrial Data

**M. Bensch[1], M. Schröder, M. Bogdan, W. Rosenstiel (Technische Informatik)**
**P. Czerner, R. Montino, G. Soberger (Elmos Semiconductor AG)**
**P. Linke, R. Schmidt (Robert Bosch GmbH)**

## Motivation

### *Problem statement*

Quality control in semiconductor manufacturing plants (FABs) is becoming increasingly complex. Production problems causing low wafer yield are often hard to find.

### *Goals*

Identify problems during wafer production
Give engineer problem insight
Robust yield prediction

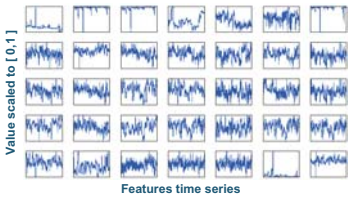
### *Methods*

Fuzzy ARTMAP neural network (FAM) [1]
Feature selection [2] (feature extraction destroys semantic information)

### *Challenges*

High data dimensionality, unbalanced classes
Process drift (non-stationarity of data due to changing machine state)
Analysed data is merely a subset of whole process

## Data




A subset of 35 electrical measurements from wafer is shown. Two classes signify "high yield" and "low yield". Ratio of training to test data is 3:1 (in chronological order). Independent test set is vital for feature selection [4].

| | Method | Features | Samples | "Low Yield" |
|---|---|---|---|---|
| Data Set 1 | Feature sel. | 61 | 3000 | 57% |
| | Test | 61 | 1000 | 31% |
| Data Set 2 | Feature sel. | 85 | 1000 | 48% |
| | Test | 85 | 313 | 57% |

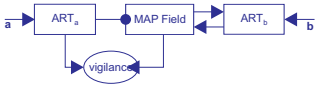
## Funding

*This work was done as part of the work package "Online Tool Controlling" (OTC), funded by the BMBF (Federal Ministry of Education and Research) and supported by the project partners Robert Bosch GmbH and Elmos Semiconductor AG. OTC is part of the "Overall Equipment Efficiency" project, also funded by the BMBF.*

## Methods

### *Yield prediction*

FAM is a suitable classifier for wafer data

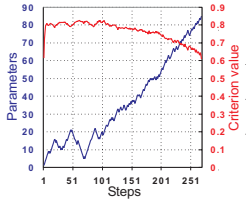



Extract Fuzzy-Rules:
"IF parameter A in [0.0476,0.524]
AND ...
THEN "Low Yield"

Few Rules — Many Rules
Easy Interpretation — High Accuracy

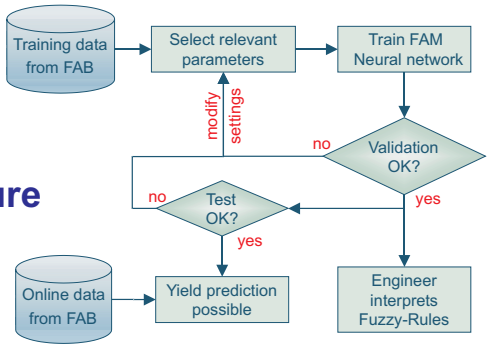
### *Feature Selection*

Sequential Forward Selection (SFS)
Sequential Forward Floating Selection (SFFS) [3]

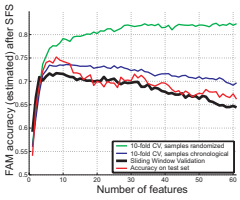



Criterion functions used:
Fuzzy ARTMAP (FAM)
k-NN (k=5, Euclidean distance)
Mahalanobis distance (M-dist)

## System Architecture




## Error Estimation

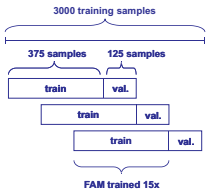



### *Problem*

Cross-validation is overoptimistic due to low within-lot variation

### *Solution*

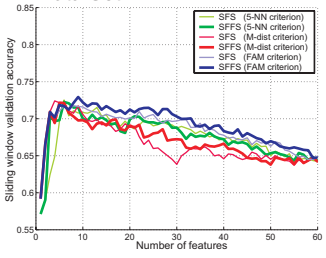
- Do not randomise data during CV
  Keep chronological order of data
- Sliding window validation (SWV)
  SWV imitates online case - data from a few lots is used to train the FAM, chronologically succeeding data is used for the test



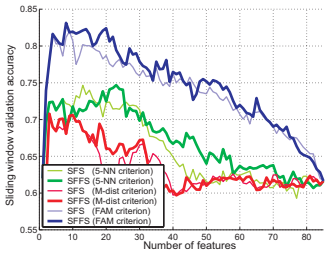

## Results

### *Sliding Window Validation*

Data Set 1

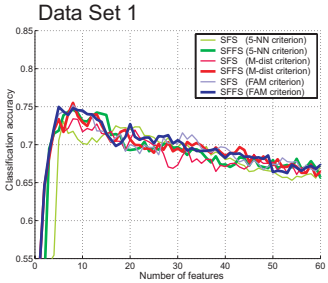



Data Set 2

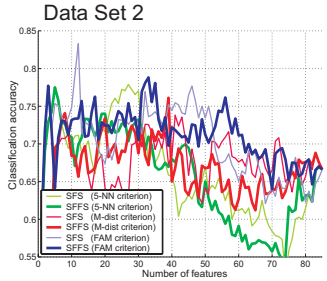



### *Yield Prediction (test)*

Data Set 1




Data Set 2




A feature subset of size 5-15 gives best yield prediction for Data Set 1.
The trend is not as clear for Data Set 2.
As expected, the FAM criterion function delivers the best results.

## Discussion

- SFFS is not superior to SFS when predicting wafer yield
- A small feature subset can model a production problem well
- Yield prediction improves after feature selection

A real production problem was solved by combining
- Feature selection
- Fuzzy Rules extracted from FAM
- Engineer's problem solving competence

### *Further Challenges*

- Visualisation of Fuzzy Rules
- More datasets needed to prove generality of results and allow online yield prediction

References:
[1] G.A. Carpenter and S. Grossberg. Fuzzy ARTMAP: A neural network architecture for incremental supervised learning of analog multidimensional maps. IEEE Transactions on Neural Networks, 1992.
[2] Isabelle Guyon and André Elisseeff. An introduction to variable and feature selection. Journal of Machine Learning Research, 2003.
[3] P. Pudil, F.J. Ferri, J. Novovicova, and J.V. Kittler. Floating search methods for feature selection with nonmonotonic criterion functions. In ICPR94, pages 279-283, 1994.
[4] Juha Reunanen. Overfitting in making comparisons between variable selection methods. Journal of Machine Learning Research, 2003.

[1] Corresponding author: bensch@informatik.uni-tuebingen.de

ELMOS BOSCH NeuroTeam
SPONSORED BY THE Federal Ministry of Education and Research