







Application of Event Based Decision Tree and Ensemble of Data Driven Methods for Maintenance Action Recommendation





James K. Kimotho, Christoph Sondermann-Woelke,
Tobias Meyer and Walter Sestro


Agenda


-  Data Preparation
-  Evolution of methods
 - Parametric based ensemble of data driven methods
 - Event based decision tree and support vector machines (SVM)
 - Event based decision tree and ensemble of data driven methods
-  Results
-  Conclusions


Data Preparation

-  Data files split into smaller files of 5000 data instances each
 - For easier handling
 - Occupies less memory

-  Processing done in Matlab environment
 - Data converted to *.mat files

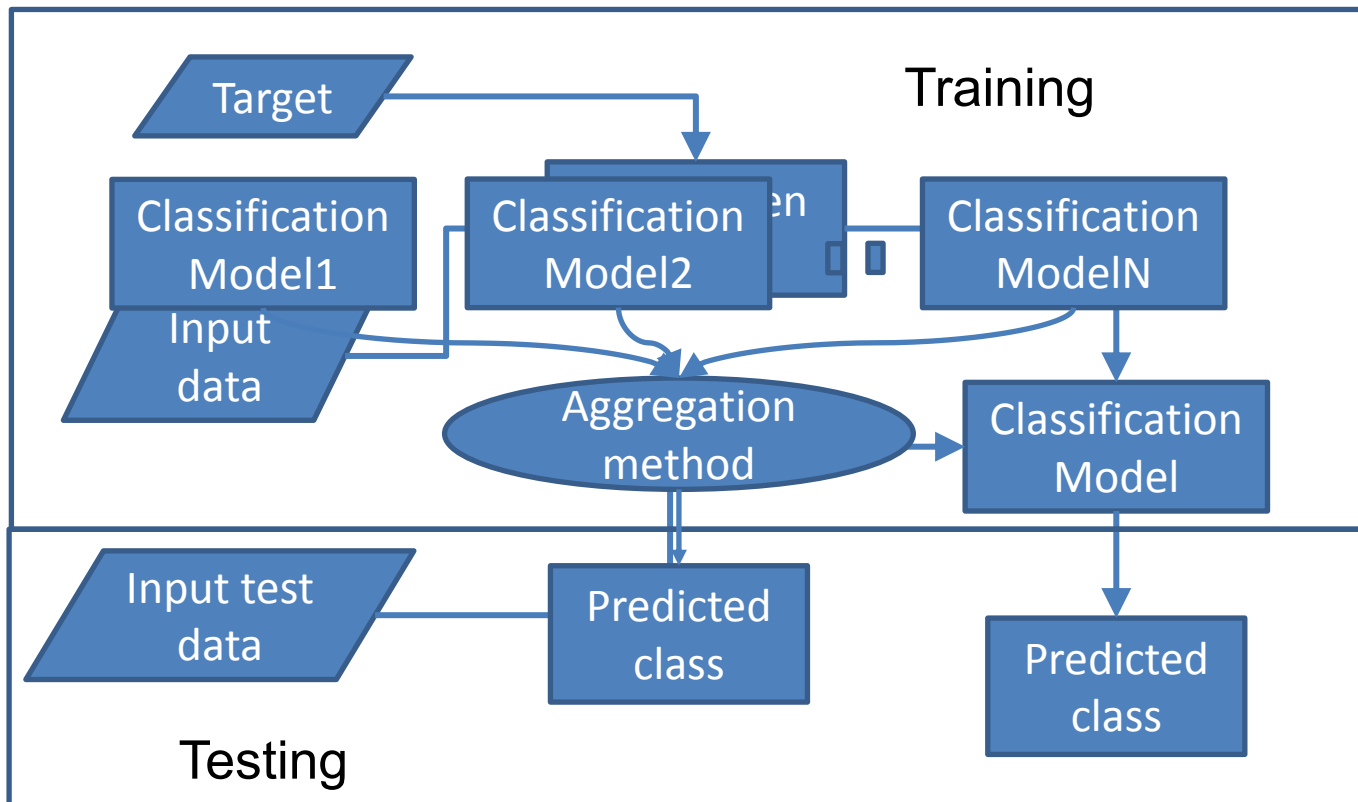
-  Separation of confirmed problems and nuisance data

-  Removal of constant parameters, repeated parameters and implausible values

-  Isolation of cases with missing information
 - processed separately

Evolution of Methods

- Parametric based ensemble of data driven methods
- Data driven methods: based on machine learning approach
- Ensemble: aggregation of results from several methods



Evolution of Methods

Parametric based ensemble of data driven methods

State of the art data driven methods:

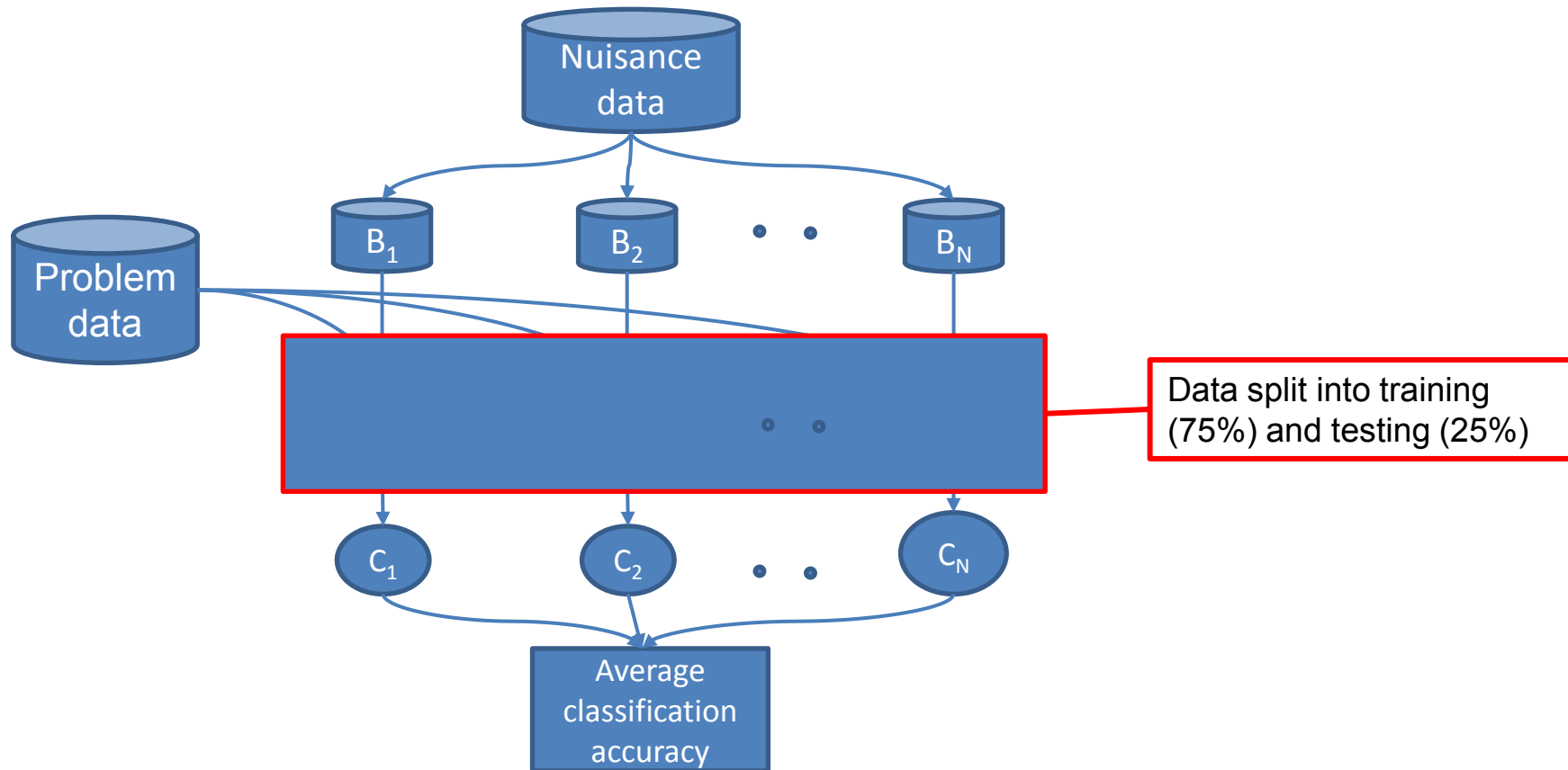
- K-Nearest Neighbors (kNN)
- Artificial Neural Networks (ANN)
- Classification and regression trees (CART)
- Bagged trees (BT)
- Random forests (RF)
- Support vector machines (SVM)

Data usage:

- Training
 - Input: parametric data
 - Target: problem labels
- Testing
 - Input: parametric data
 - Output: predicted problem labels

Evolution of Methods

- Parametric based ensemble of data driven methods
- Training approach for each algorithm
 - Batch training using 40 nuisance cases per batch



Evolution of Methods

Parametric based ensemble of data driven methods

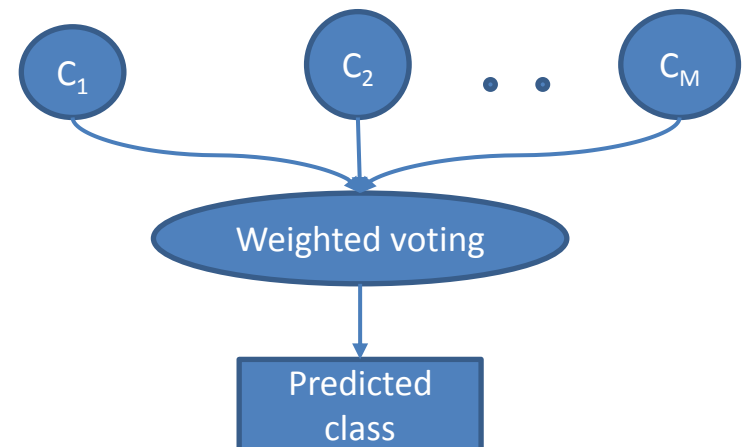
Performance Evaluation

| Algorithm | Average accuracy |
|-----------------|------------------|
| kNN | 89% |
| ANN | 88% |
| CART | 83% |
| BT | 93% |
| RF | 94% |
| SVM | 90% |
| Ensemble | 95% |

Training accuracy: quite impressive 😊

Ensemble of algorithms

- Aggregation of results from several algorithms
- Weighted voting to obtain final class



Score: 21 😞

Evolution of Methods

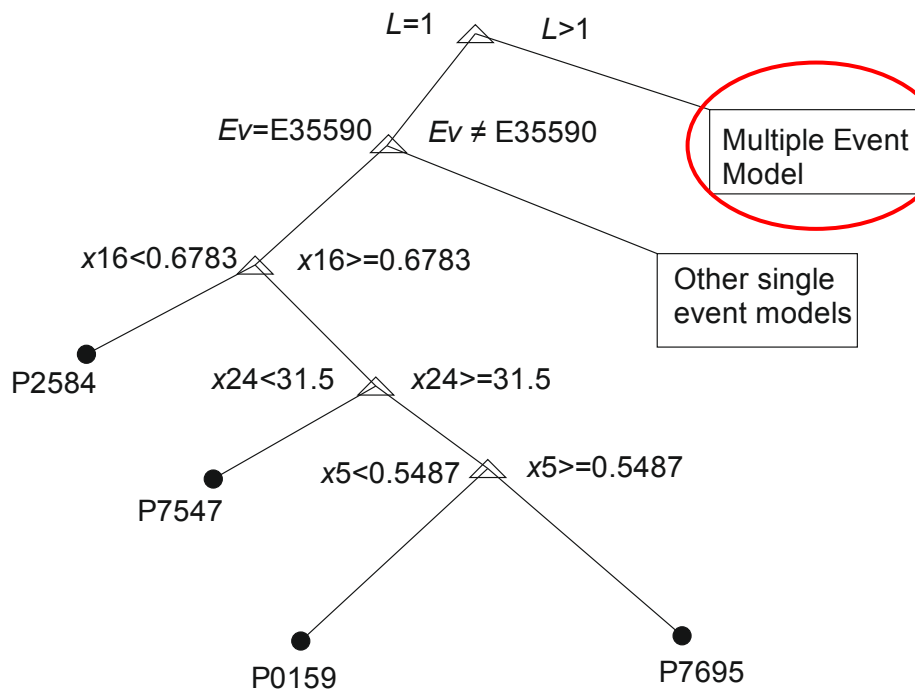
Event based decision tree and SVM

- How can we use the event code?
- Event code indicates:
 - System or subsystem that measurements are taken from
 - Reason why the code was generated
- Two categories of cases were identified:
 - Cases with single events
 - Cases with multiple events
- Decision tree constructed based on event codes
- Parametric data used to derive decision rules for events appearing in multiple problems.
- For cases with multiple events, SVM was used to build a model for each event
 - SVM selected since it is robust

Evolution of Methods

Event based decision tree and SVM

Construction of event based decision tree








- SVM with optimally tuned parameters
 - Parameters tuned using particle swarm optimization (PSO)
- A classification model was built for each event
 - 165 models built
- Input: parametric data

Score: 48

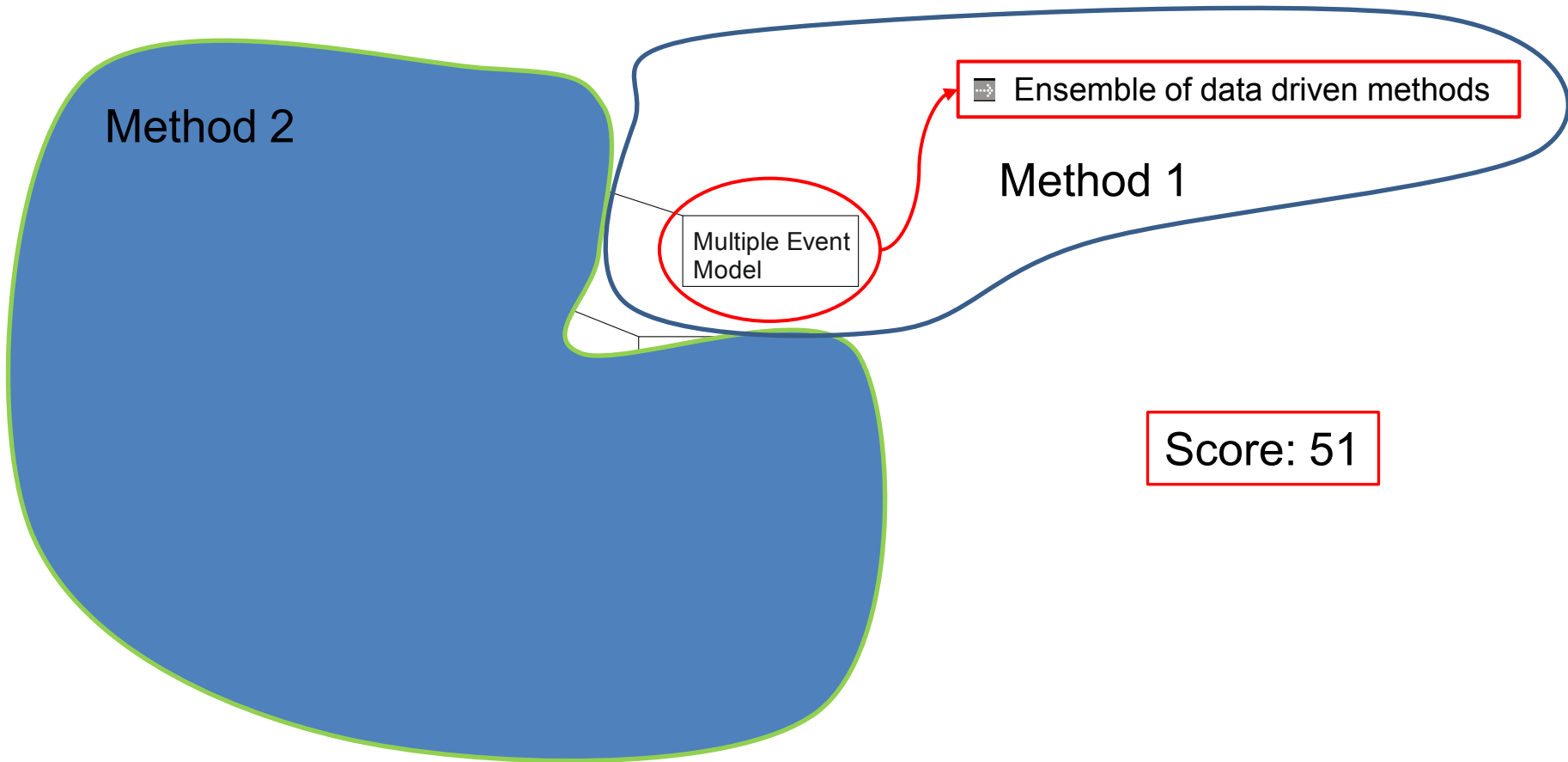
L is No. Of unique events per case

Evolution of Methods

-  Event based and ensemble of data driven methods
-  Ensemble method employed to improve classification
-  Methods employed:
 - SVM with optimally tuned parameters
 - Bagged trees
 - Random forests
-  165 models of each algorithm were built
-  Final classification obtained through majority voting

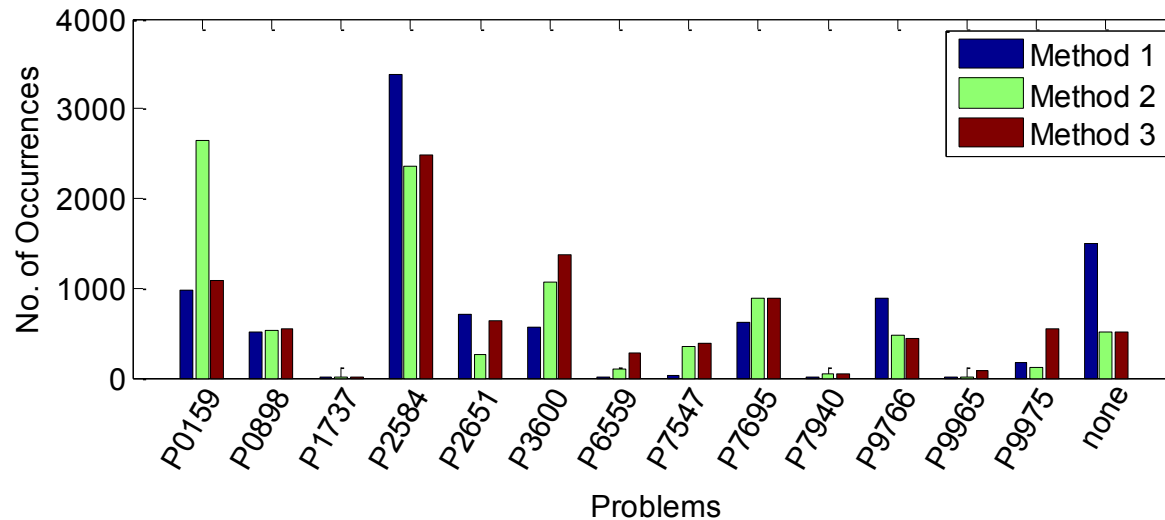
Evolution of Methods

- Event based and ensemble of data driven Methods
- Construction of event based decision tree for event E35590



Results

Distribution of classification results of the test data



Last two methods: nuisance data considered only for single events

Performance analysis

| | No. of outputs | No. of incorrect outputs | No. of nuisance outputs | Score |
|----------|----------------|--------------------------|-------------------------|-------|
| Method 1 | 303 | 133 | 149 | 21 |
| Method 2 | 332 | 122 | 162 | 48 |
| Method 3 | 331 | 121 | 159 | 51 |



Conclusions

- Methodologies for maintenance action recommendation based on event codes and parametric data have been presented
- Combination of event based decision tree and ensemble of data driven methods improves classification
- Methodologies presented led to our team being ranked position three
- Method can be improved if more information on the equipment and parametric data is revealed



UNIVERSITÄT PADERBORN
Die Universität der Informationsgesellschaft

Fakultät für Maschinenbau
Lehrstuhl für Mechatronik und Dynamik
Prof. Dr.-Ing. habil. Walter Sestro

Thank you for your kind attention!

James K. Kimotho
james.kuria.kimotho@uni-paderborn.de
Tel. 05251/60-1819
Fax 05251/60-1803

Mechatronik und Dynamik
Universität Paderborn
Pohlweg 47-49
33098 Paderborn