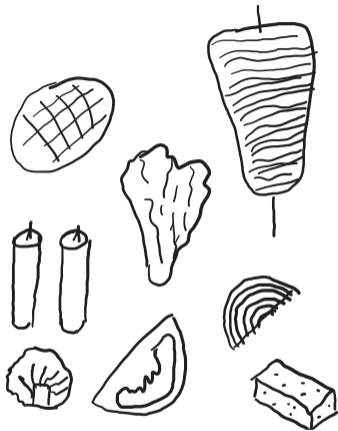
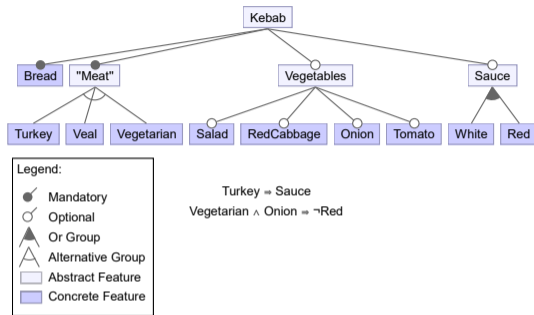
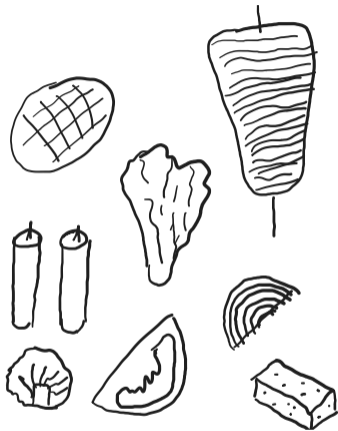


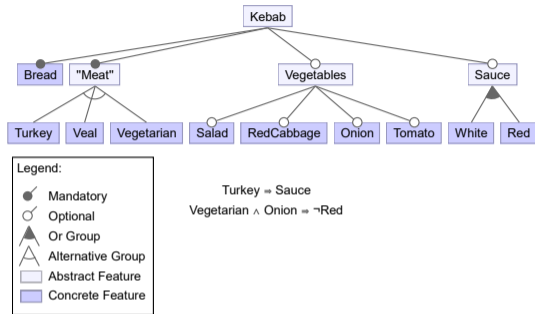
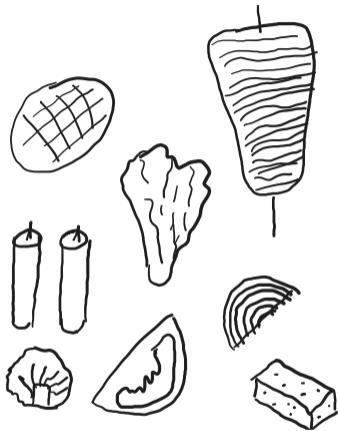


## Engineering and Selecting Features for Configuration Models

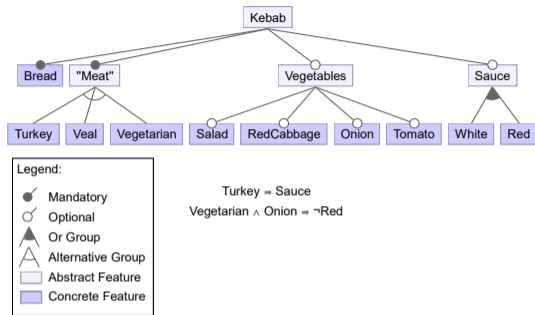
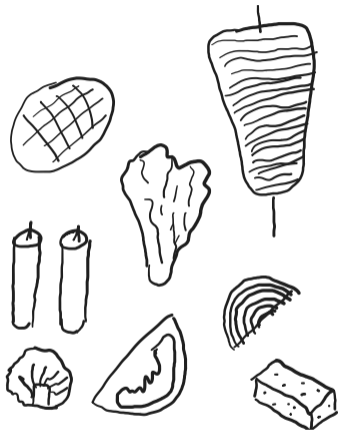
FOSD'26 | Raphael Dunkel, Elias Kuitert, Chico Sundermann, Matthias Tichy, Thomas Thüm | March 27, 2026





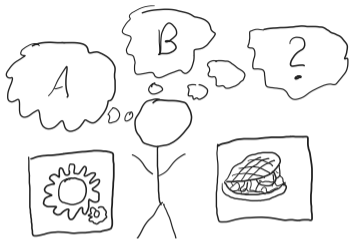


How many valid Kebabs are there?

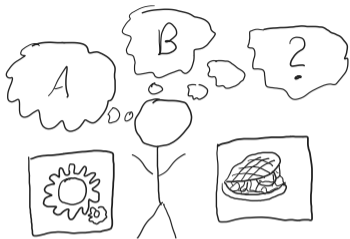


How many valid Kebabs are there?

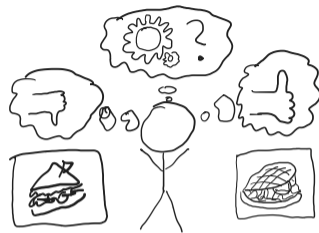




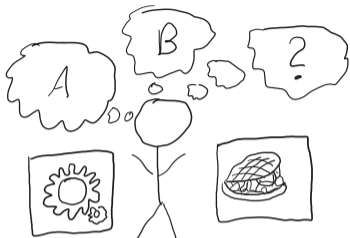
Which solver is best for my instance?



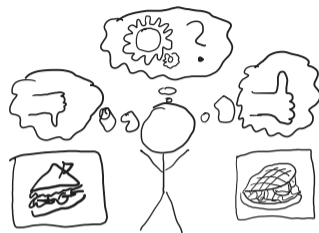
Which solver is best for my instance?



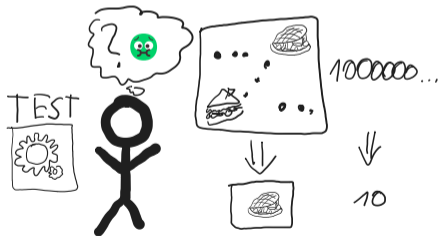
What makes configuration models complex?



Which solver is best for my instance?



What makes configuration models complex?



Which instances do I have to test?

## **Intra-Instance Machine Learning:**

## **Inter-Instance Machine Learning**

### **Intra-Instance Machine Learning:**

- On single product line

### **Inter-Instance Machine Learning**

- on multiple product lines

### **Intra-Instance Machine Learning:**

- On single product line
- (Often) configuration options as features

### **Inter-Instance Machine Learning**

- on multiple product lines
- metric-based feature set

## Intra-Instance Machine Learning:

- On single product line
- (Often) configuration options as features
- Example: performance model



[1] Performance Models  
Guo et al. in EMSE '18

## Inter-Instance Machine Learning

- on multiple product lines
- metric-based feature set
- Example: configuration model maintainability

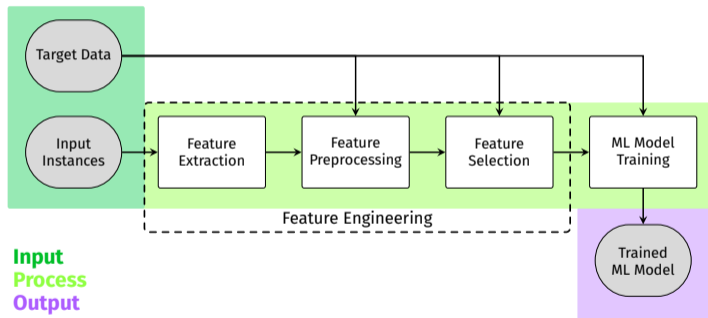


[2] Configuration Model Maintainability  
Silva et al. at SPLC '21

[1] Guo et al., "Data-Efficient Performance Learning for Configurable Systems" (2018, Kluwer Academic Publishers)  
[2] Silva et al., "A Machine Learning Model To Classify the Feature Model Maintainability" (2021, ACM)

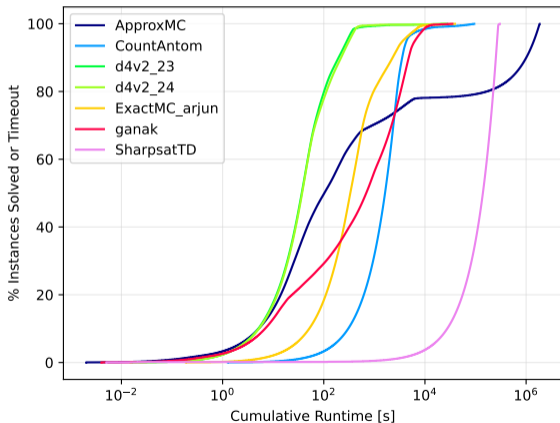
# Contributions

## 1. ML Framework FE4FEMO



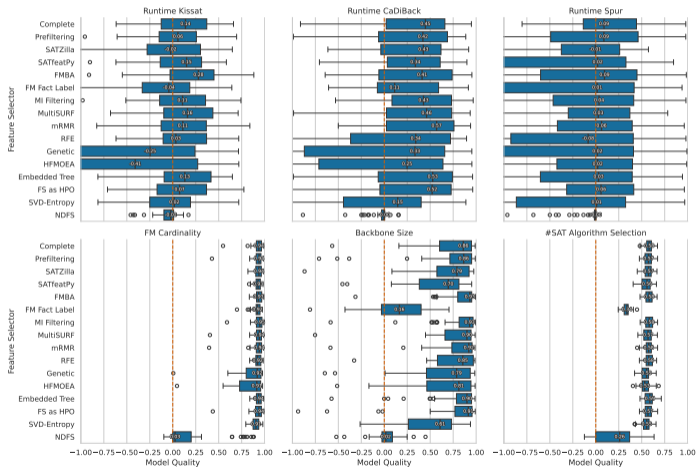
# Contributions

1. ML Framework FE4FEMO
2. New Data Sets for ML on Configuration Models
  - Target Data
  - Extracted Features



# Contributions

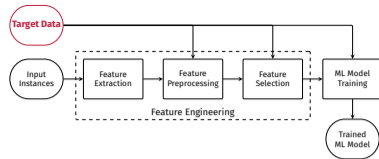
1. ML Framework FE4FEMO
2. New Data Sets for ML on Configuration Models
  - Target Data
  - Extracted Features
3. Evaluation of ML Pipeline Configurations



# Target Data

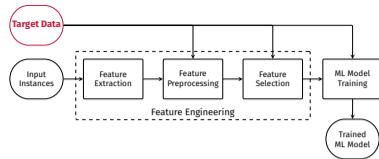


[3] FM Analysis  
Sundermann et al. in AMAI '24



[3] Sundermann et al., “On the Benefits of Knowledge Compilation for Feature-Model Analyses” (2024, Springer)

# Target Data



Category	Task	Name	Extracted
SAT	Void	KISSAT	wallclock, memory
	Core-Dead	CADIBACK	wallclock, memory, backbone size
#SAT	Feature Model Cardinality	APPROXMC	wallclock, memory, algorithm selection, #SAT value
		COUNTANOM	
		D4V2-23	
		D4V2-24	
	Uniform Random Sampling	EXACTMC_ARJUN	wallclock, memory
		GANAK	
		SHARPSATD	
		SPUR	

## [3] FM Analysis

Sundermann et al. in AMAI '24

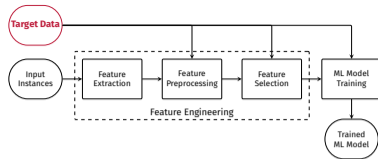
[3] Sundermann et al., "On the Benefits of Knowledge Compilation for Feature-Model Analyses" (2024, Springer)

# Target Data



## [3] FM Analysis

Sundermann et al. in AMAI '24

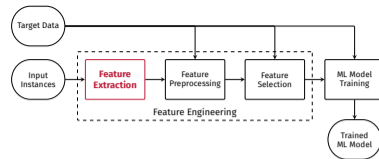


Category	Task	Name	Extracted
SAT	Void	KISSAT	wallclock, memory
	Core-Dead	CADIBACK	wallclock, memory, backbone size
#SAT	Feature Model Cardinality	APPROXMC	wallclock, memory, algorithm selection, #SAT value
		COUNTANOM	
		D4V2-23	
		D4V2-24	
	Uniform Random Sampling	EXACTMC_ARJUN	wallclock, memory
		GANAK	
		SHARPSATTD	
		SPUR	

⇒ Reusable Data Set for ML on Configuration Models

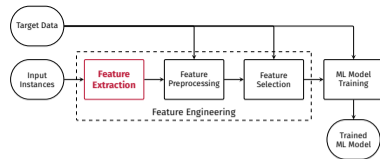
[3] Sundermann et al., “On the Benefits of Knowledge Compilation for Feature-Model Analyses” (2024, Springer)

# Feature Extraction



Domain of Origin	Feature Set	Source	#Features
Satisfiability Solving (SAT)	SATZilla	[4]	138
	SATfeatPy	[5]	352
Configuration Modeling	FMBA	[6]	28
	FM Fact Label	[7]	105
	DyMMer	[8]	38
	ESYES	[9]	14

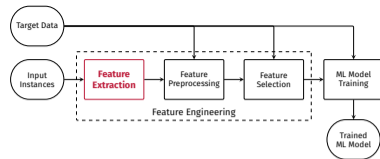
# Feature Extraction



Domain of Origin	Feature Set	Source	#Features
Satisfiability Solving (SAT)	SATZilla	[4]	138
	SATfeatPy	[5]	352
Configuration Modeling	FMBA	[6]	28
	FM Fact Label	[7]	105
	DyMMer	[8]	38
	ESYES	[9]	14

- Feature Groups
- Bug Fixes
- Optimizations: Caching, ...

# Feature Extraction

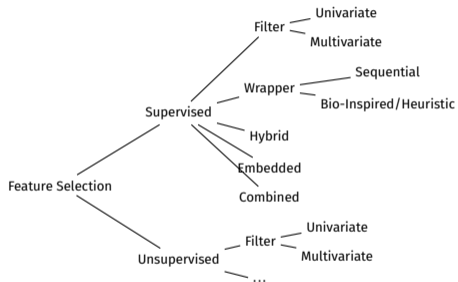


- Feature Groups
- Bug Fixes
- Optimizations: Caching, ...

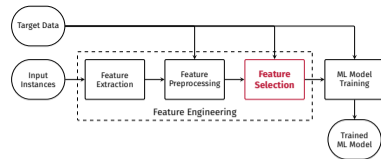
⇒ Reusable Feature Values

Domain of Origin	Feature Set	Source	#Features
Satisfiability Solving (SAT)	SATZilla	[4]	138
	SATfeatPy	[5]	352
Configuration Modeling	FMBA	[6]	28
	FM Fact Label	[7]	105
	DyMMer	[8]	38
	ESYES	[9]	14

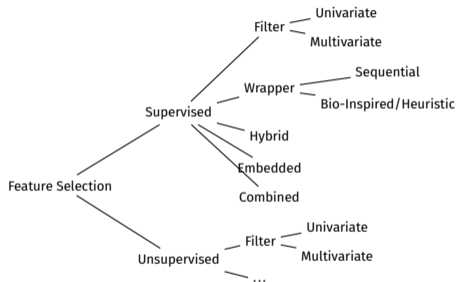
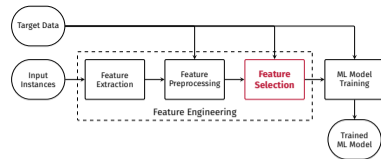
# Feature Selection



Adapted from [10, 11]



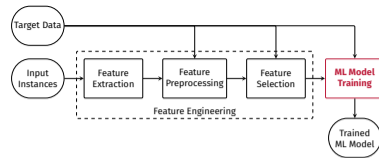
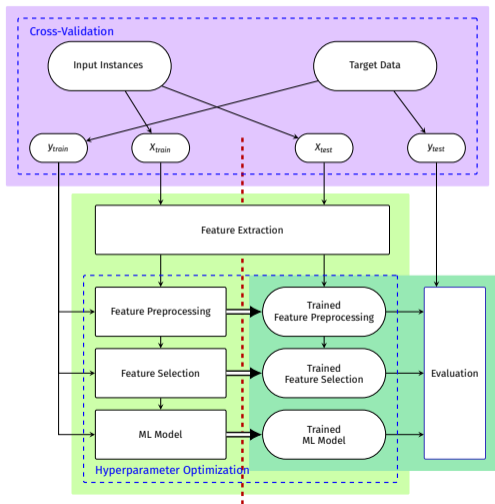
# Feature Selection



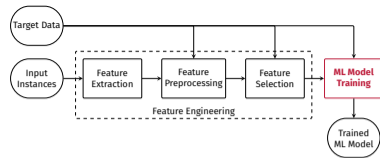
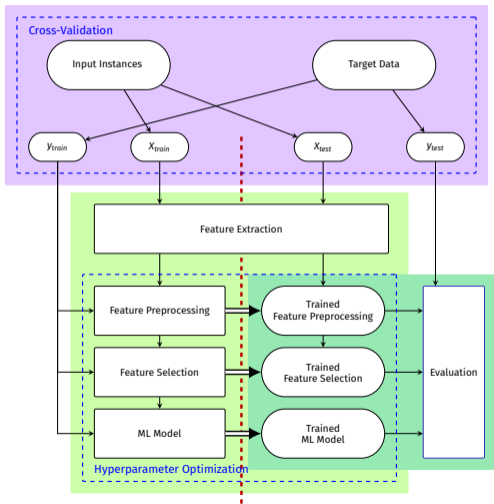
Adapted from [10, 11]

Category		Name	
Supervised	Filter	Univariate	Mutual Information Filtering
		Multivariate	MultiSURF
	Wrapper		mRMR
		Sequential	Recursive Feature Elimination
		Bio-Inspired / Heuristic	Genetic Algorithm
		Hybrid	HFMOEA
Unsupervised	Filter	Embedded	Embedded Tree-based FS
		Combined	Feature Selection as HPO
	Filter	Univariate	SVD-Entropy Filtering
		Multivariate	NDFS

# ML Model Training

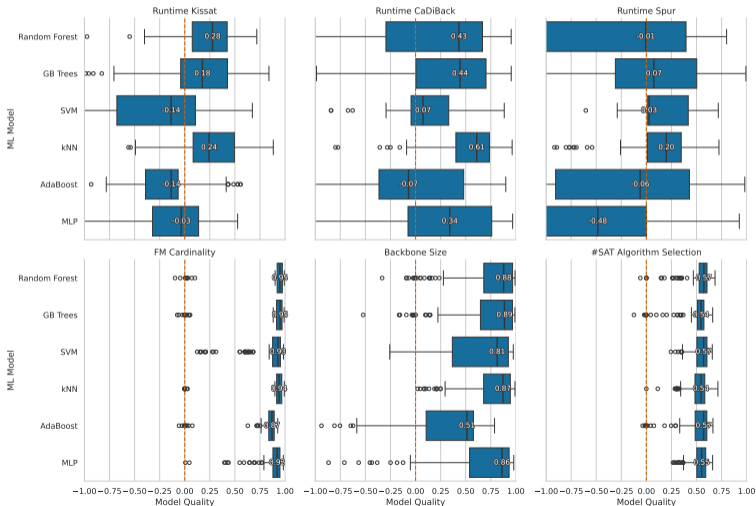
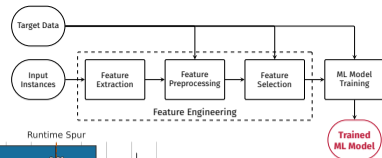


# ML Model Training

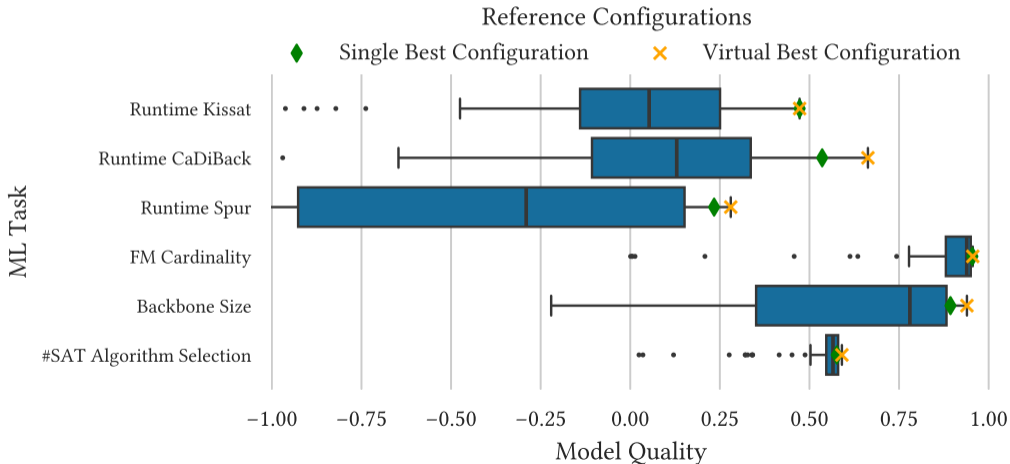
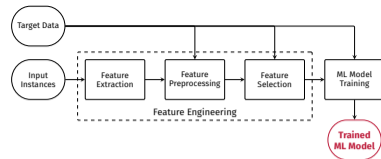


Model Name	Approaches
Random Forest	tree, ensemble, bootstrap aggregation
Gradient-Boosted Trees	tree, ensemble, boosting
Support Vector Machine	linear
<i>k</i> -Nearest Neighbors	clustering
AdaBoost	ensemble, boosting, (tree)
Multilayer Perceptron	neural network

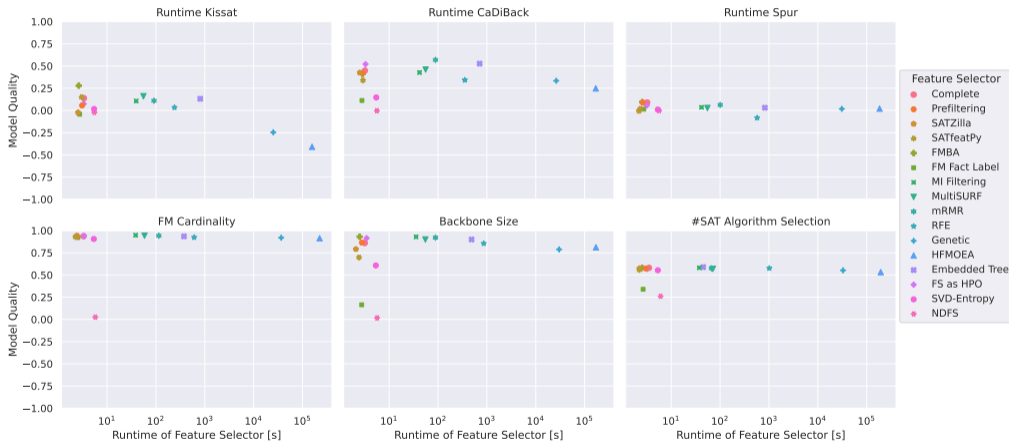
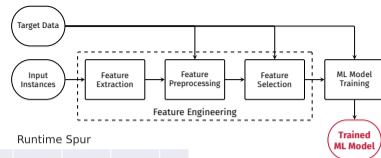
# Evaluation – Performance



# Evaluation – Potential



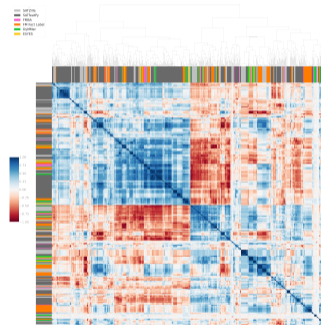
# Evaluation – Time-Performance Trade-Off



# Future Work & Overview

## Future Work

- Feature space analysis
- Complexity measure for configuration models
- Feature-based similarity metric
- Benchmark composition analysis through clustering
- Guided generation of similar artificial models
- ...



# Future Work & Overview

## Future Work

- Feature space analysis
- Complexity measure for configuration models
- Feature-based similarity metric
- Benchmark composition analysis through clustering
- Guided generation of similar artificial models
- ...

## Overview

- ML framework FE4FEMO
- Reusable data sets
- Pipeline configuration impact analysis



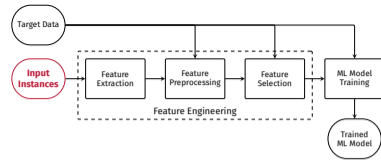
# Bibliography I

- [1] Jianmei Guo et al. “Data-Efficient Performance Learning for Configurable Systems”. June 2018
- [2] Publio Silva, Carla I. M. Bezerra, and Ivan Machado. “A Machine Learning Model To Classify the Feature Model Maintainability”. 2021
- [3] Chico Sundermann et al. “On the Benefits of Knowledge Compilation for Feature-Model Analyses”. Oct. 2024
- [4] Hadar Shavit and Holger H. Hoos. “Revisiting SATZilla Features in 2024”. 2024
- [5] Marco Dalla et al. “SAT Feature Analysis for Machine Learning Classification Tasks”. 2023
- [6] Chico Sundermann et al. “Evaluating State-of-the-Art #SAT Solvers on Industrial Configuration Spaces”. Mar. 2023
- [7] Jos’e Miguel Horcas et al. “FM Fact Label”. 2025
- [8] Anderson G. Uchôa et al. “DyMMer-NFP: Modeling Non-functional Properties and Multiple Context Adaptation Scenarios in Software Product Lines”. 2017
- [9] Sascha El-Sharkawy, Nozomi Yamagishi-Eichler, and Klaus Schmid. “Metrics for Analyzing Variability and its Implementation in Software Product Lines: A Systematic Literature Review”. 2019

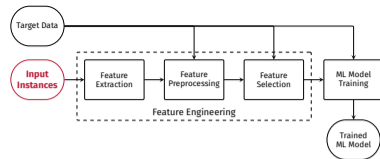
## Bibliography II

- [10] Yvan Saeys, Iñaki Inza, and Pedro Larrañaga. “A Review of Feature Selection Techniques in Bioinformatics”. 2007
- [11] Saúl Solorio-Fernández, J. Ariel Carrasco-Ochoa, and José Fco. Martínez-Trinidad. “A Review of Unsupervised Feature Selection Methods”. 2020
- [12] Chico Sundermann et al. “Collecting Feature Models from the Literature: A Comprehensive Dataset for Benchmarking”. Sept. 2024

# Input Instances



# Input Instances



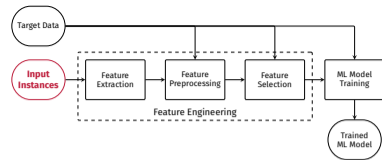
## [12] Feature-Model Benchmark Sundermann et al. in SPLC '24

[12] Sundermann et al., "Collecting Feature Models from the Literature: A Comprehensive Dataset for Benchmarking" (2024, ACM)

# Input Instances



- biggest collection of feature models
- 2334 usable feature models
- 101 – 80,258 features
- various domains
- contains feature model histories



[12] Feature-Model Benchmark  
Sundermann et al. in SPLC '24

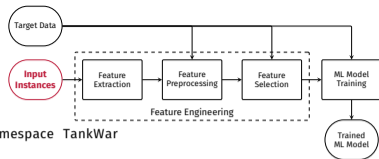
[12] Sundermann et al., “Collecting Feature Models from the Literature: A Comprehensive Dataset for Benchmarking” (2024, ACM)

# Input Instances



[12] Feature-Model Benchmark  
Sundermann et al. in SPLC '24

- biggest collection of feature models
- 2334 usable feature models
- 101 – 80,258 features
- various domains
- contains feature model histories
- UVL



namespace TankWar

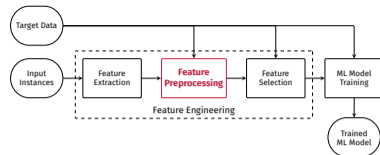
```

features
  TankWar {abstract true}
    mandatory
      Tank_Player1 {abstract true}
        alternative
          Black_P1
          Red_P1
          Red_Tier1_P1
          Yellow_Tier1_P1
          Blue_Tier2_P1
          Green_Tier2_P1
      Tank_Player2 {abstract true}
        alternative
          Black_P2
          Red_P2
          Red_Tier1_P2
          Yellow_Tier1_P2
          Blue_Tier2_P2
          Green_Tier2_P2
    mov {abstract true}
      alternative
        mov_0
        mov_1
        mov_2
    
```

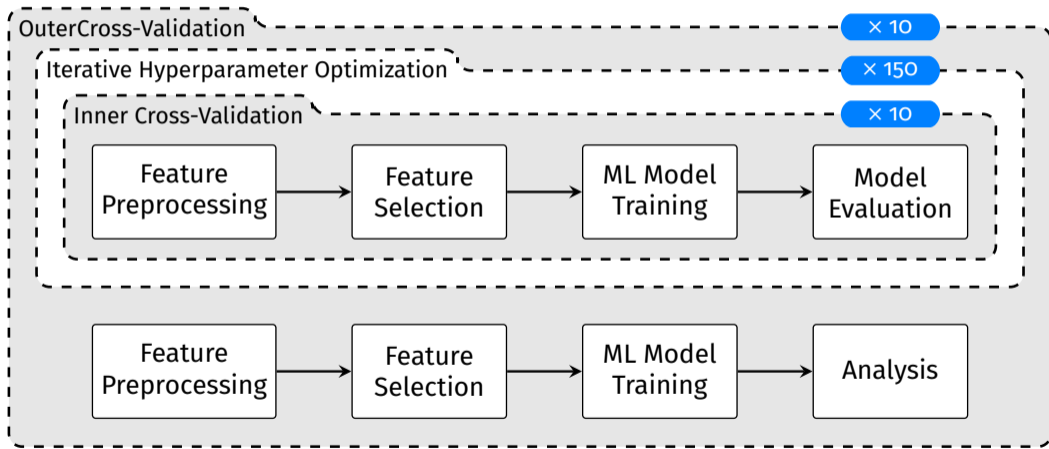
[...]

[12] Sundermann et al., “Collecting Feature Models from the Literature: A Comprehensive Dataset for Benchmarking” (2024, ACM)

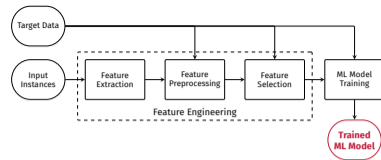
# Feature Preprocessing



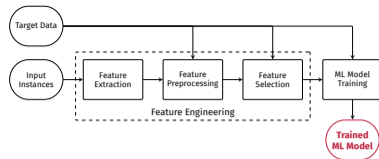
# Cross-Validation Strategy



# Evaluation – Research Questions

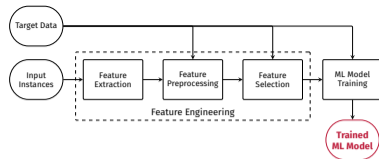


# Evaluation – Research Questions



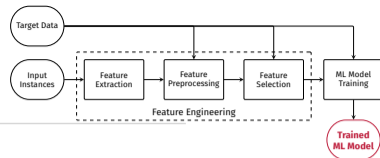
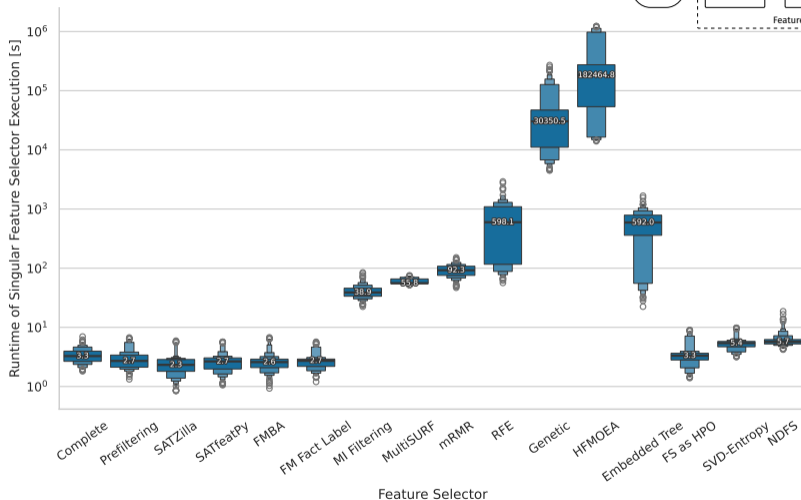
1. How well do different **ML pipeline configurations**, especially **feature selection strategies**, perform for exemplary ML tasks on feature model analysis with regards to ...
  1. ...ML feature set **reduction**?
  2. ...feature selection **stability**?
  3. ...ML model **prediction quality**?
  4. What is the **trade-off** between all performance metrics?
2. How difficult are different **ML pipeline configurations**, especially **feature selection strategies**, to compute for exemplary ML tasks on feature model analysis with regards to ...
  1. ...feature selection algorithm **runtime**?
  2. ...active ML feature subset **computation time**?
  3. What is the **trade-off** between runtime, ML feature subset computation time, and ML model quality?
3. How is the combined **feature space** structured in terms of ...
  1. ...ML **feature correlation**?
  2. ...ML **feature importance**?

# Evaluation – Research Questions



1. How well do different **ML pipeline configurations**, especially **feature selection strategies**, perform for exemplary ML tasks on feature model analysis with regards to ...
  1. ...ML feature set **reduction**?
  2. ...feature selection **stability**?
  3. ...ML model **prediction quality**?
  4. What is the **trade-off** between all performance metrics?
2. How difficult are different **ML pipeline configurations**, especially **feature selection strategies**, to compute for exemplary ML tasks on feature model analysis with regards to ...
  1. ...feature selection algorithm **runtime**?
  2. ...active ML feature subset **computation time**?
  3. What is the **trade-off** between runtime, ML feature subset computation time, and ML model quality?
3. How is the combined **feature space** structured in terms of ...
  1. ...ML **feature correlation**?
  2. ...ML **feature importance**?

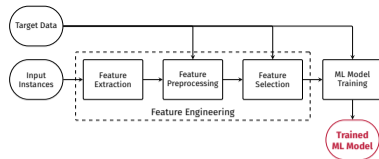
# Evaluation – RQ 2.1: Computational Effort



# Evaluation – Threats to Validity

## Internal Validity:

- Target Data Generation
  - Measurement errors
  - Right censoring
- Feature Extraction
  - Implementation errors
- Feature Preprocessing
  - Prefiltering can remove good features
  - Non-optimal imputation assumption
  - Downcasting
- Feature Selection
  - Implementation errors
- ML Model Training
  - Overfitting
  - Performance metrics



## External Validity:

- Target Data Generation
  - Different Solvers could have different results
- Chosen Input Instances
  - Representative?
  - Groups in data set
  - Feature model size
- Selected ML Pipeline Configurations
  - Subset representative?
- No Established Evaluation Standards

# Applications and Improvements

## Improving Feature Extraction:

- Graph-based Features
- Solution Space Metrics
- Autoencoders
- Convolutional Neural Networks

## Improving the Pipeline:

- Feature Transformations
- Imputation with MAR
- Feature Importance Analysis

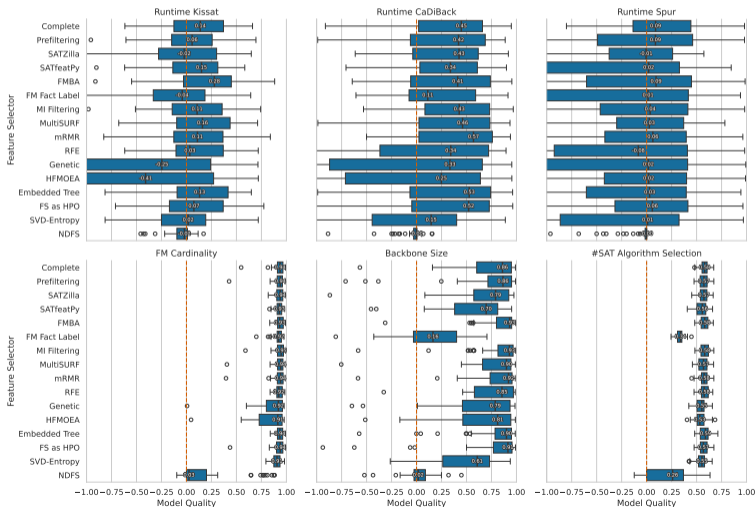
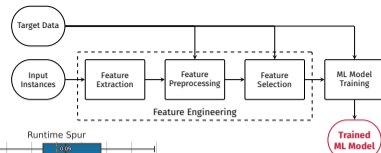
## Analyzing FM Benchmarks:

- Clustering Analysis
- Benchmark Minimization
- Synthetic Benchmark Expansion

## Solving Other Problems:

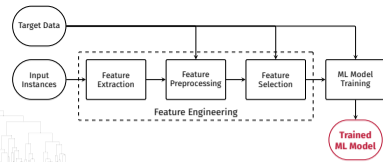
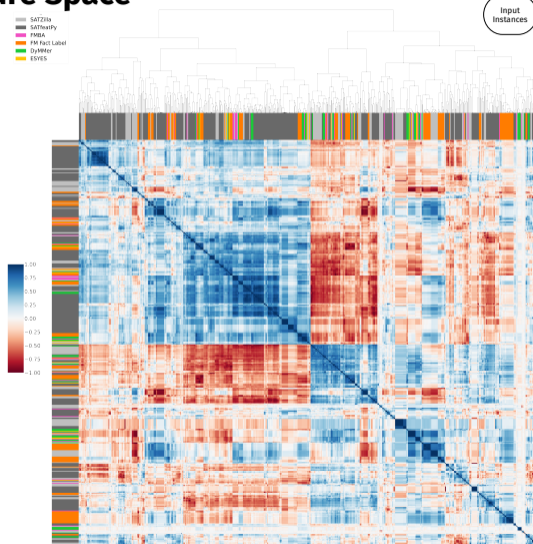
- #SAT Algorithm Selection
- FM Similarity
- FM Complexity Metrics

# Evaluation – Performance



# Evaluation – Feature Space

- SATZ06
- SATZ07
- PMBA
- PM Fact Label
- DyMMer
- ESSES



# Engineering and Selecting Features for Configuration Models

**1. Motivation**

**2. Contributions**

**3. fe4femo**

**4. Evaluation**

**5. Outlook**

**6. Appendix**