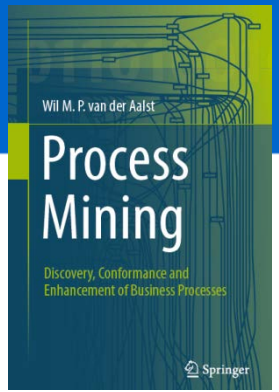
Process Mining: Data Science in Action

Evaluating Mining Results



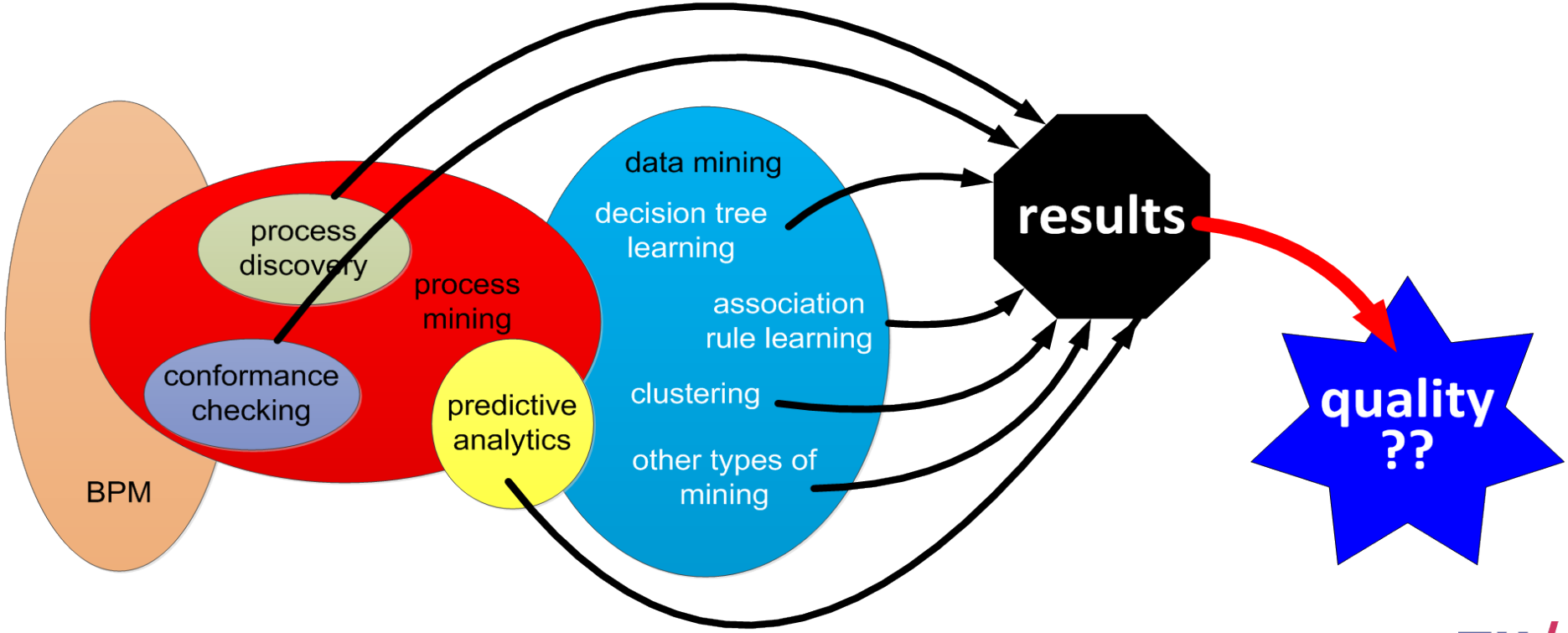
prof.dr.ir. Wil van der Aalst
www.processmining.org

TU/e Technische Universiteit Eindhoven University of Technology

Where innovation starts

©Wil van der Aalst & TU/e (use only with permission & acknowledgements)

Evaluating (data/process) mining results

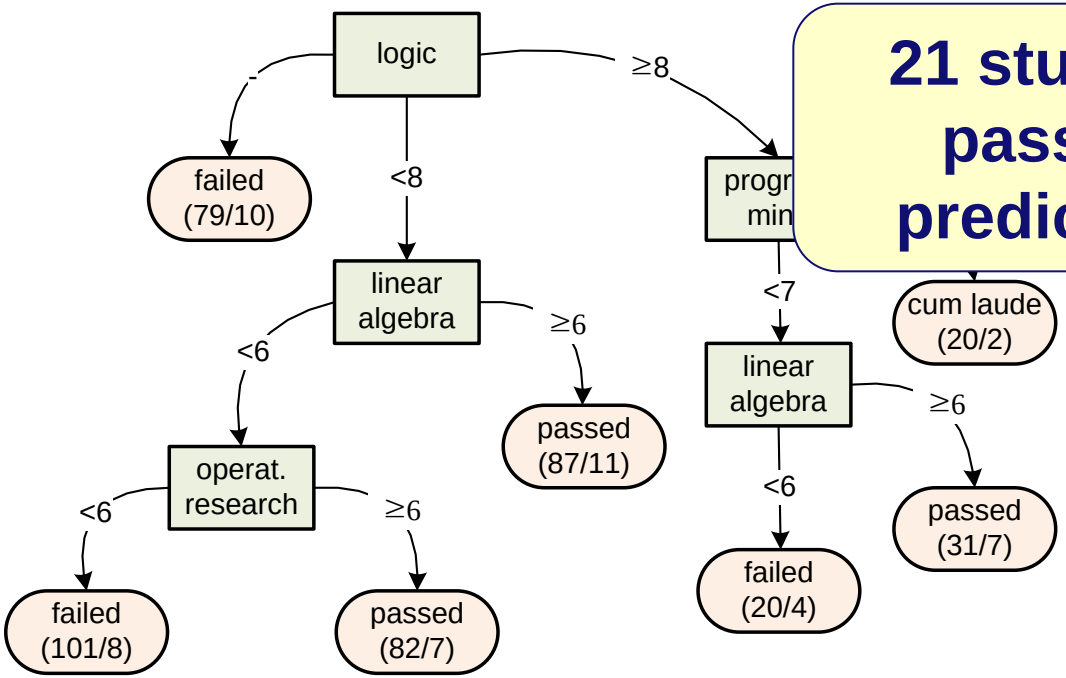


©Wil van der Aalst & TU/e (use only with permission & acknowledgements)

TU/e

Confusion matrix and related measures

Confusion matrix for decision tree



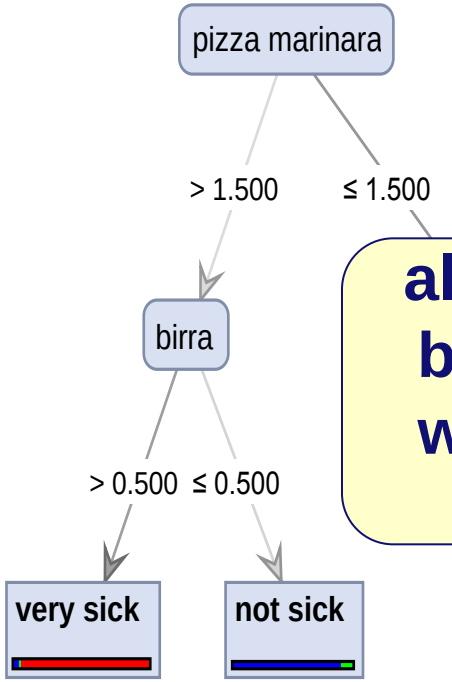
21 students that passed were predicted to fail

		predicted class		
		failed	passed	cum laude
actual class	fa..d	178	22	0
	passed	21	175	2
	cum laude	1	3	18

©Wil van der Aalst & TU/e (use only with permission & acknowledgements)

TU/e

Another example



accuracy: 93.26%

	true
pred. not sick	419
pred. nauseous	0
pred. very sick	24
class recall	99.4

none of the parties were predicted to become nauseous

all 470 parties that became very sick were predicted to become sick

		predicted class: not sick	predicted class: nauseous	predicted class: very sick
actual class	not sick	4193	0	24
	nauseous	307	0	6
	very sick	0	0	470

©Wil van der Aalst & TU/e (use only with permission & acknowledgements)

TU/e

Confusion matrix for binary classification

- **True Positives (TP): positive instances predicted to be positive .**
- **True Negatives (TN): negative instances predicted to be negative.**
- **False Positives (FP): negative instances predicted to be positive.**
- **False Negatives (FN): positive instances predicted to be negative.**

		predicted class	
		+	-
actual class	+	TP	FN
	-	FP	TN

©Wil van der Aalst & TU/e (use only with permission & acknowledgements)

- **True Positives (TP): sick customers predicted to be sick.**
- **True Negatives (TN): non-sick customers predicted to be not sick.**
- **False Positives (FP): non-sick customers predicted to be sick.**
- **False Negatives (FN): sick customers predicted to be not sick.**

		predicted class	
		+	-
actual class	+	TP	FN
	-	FP	TN

Quality measures
(based on confusion matrix)

- **error** = (FP+FN)/K
- **accuracy** = (TP+TN)/K
- **precision** = TP/P' = TP/(TP+FP)
- **recall** = TP/P = TP/(TP+FN)
- **F1-score** =
  (2 x precision x recall)/(precision + recall)
  (harmonic mean of precision and recall)

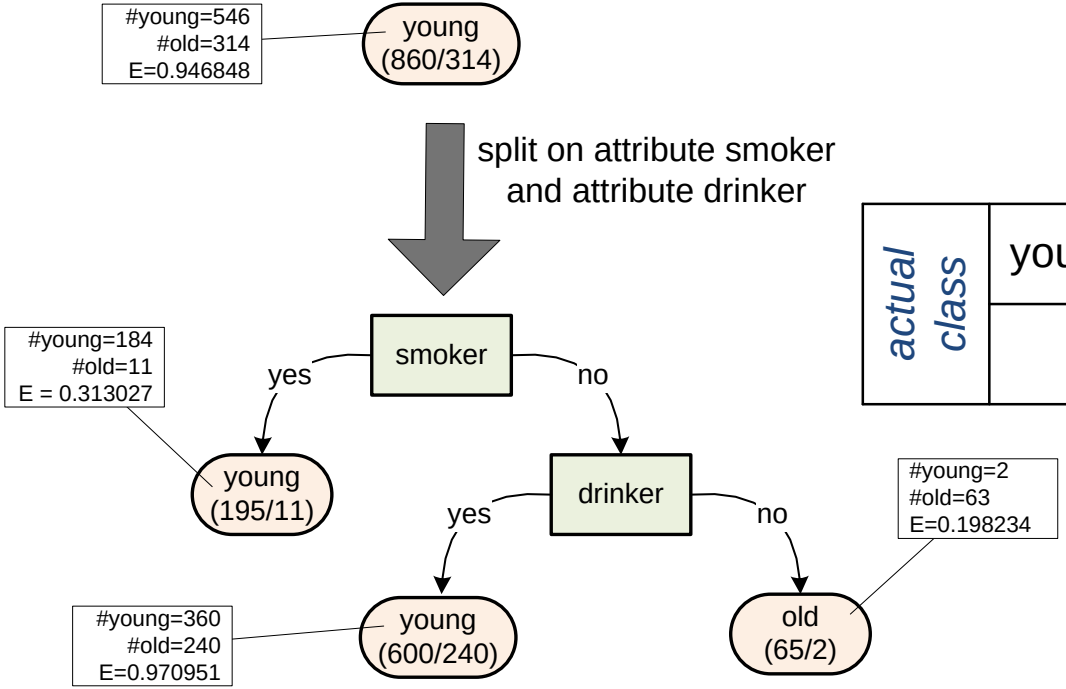
		predicted class +	predicted class -	
actual class	+	TP	FN	P = TP+FN
actual class	-	FP	TN	N = FP+TN
		P' = TP+FP	N' = FN+TN	K = TP+FN+FP+TN

©Wil van der Aalst & TU/e (use only with permission & acknowledgements)

TU/e

Question: Compute precision, recall, and the F1-score before and after splitting

young = positive
old = negative



		predicted class: young	predicted class: old
actual class	young	546	0
	old	314	0

		predicted class: young	predicted class: old
actual class	young	544	2
	old	251	63

©Wil van der Aalst & TU/e (use only with permission & acknowledgements)

TU/e

Answer

precision = 546/(546+314) = 0.635
recall = 546/(546+0) = 1.000
F1-score = 0.777

		predicted class: young	predicted class: old
actual class	young	546	0
	old	314	0

		predicted class: young	predicted class: old
actual class	young	544	2
	old	251	63

precision = 544/(544+251) = 0.684
recall = 544/(544+2) = 0.996
F1-score = 0.811

©Wil van der Aalst & TU/e (use only with permission & acknowledgements)

TU/e

Cross-validation

Consider your 10 best friends

You can create a decision tree that accurately predicts the length of a friend based on his/her birth date and eye color.

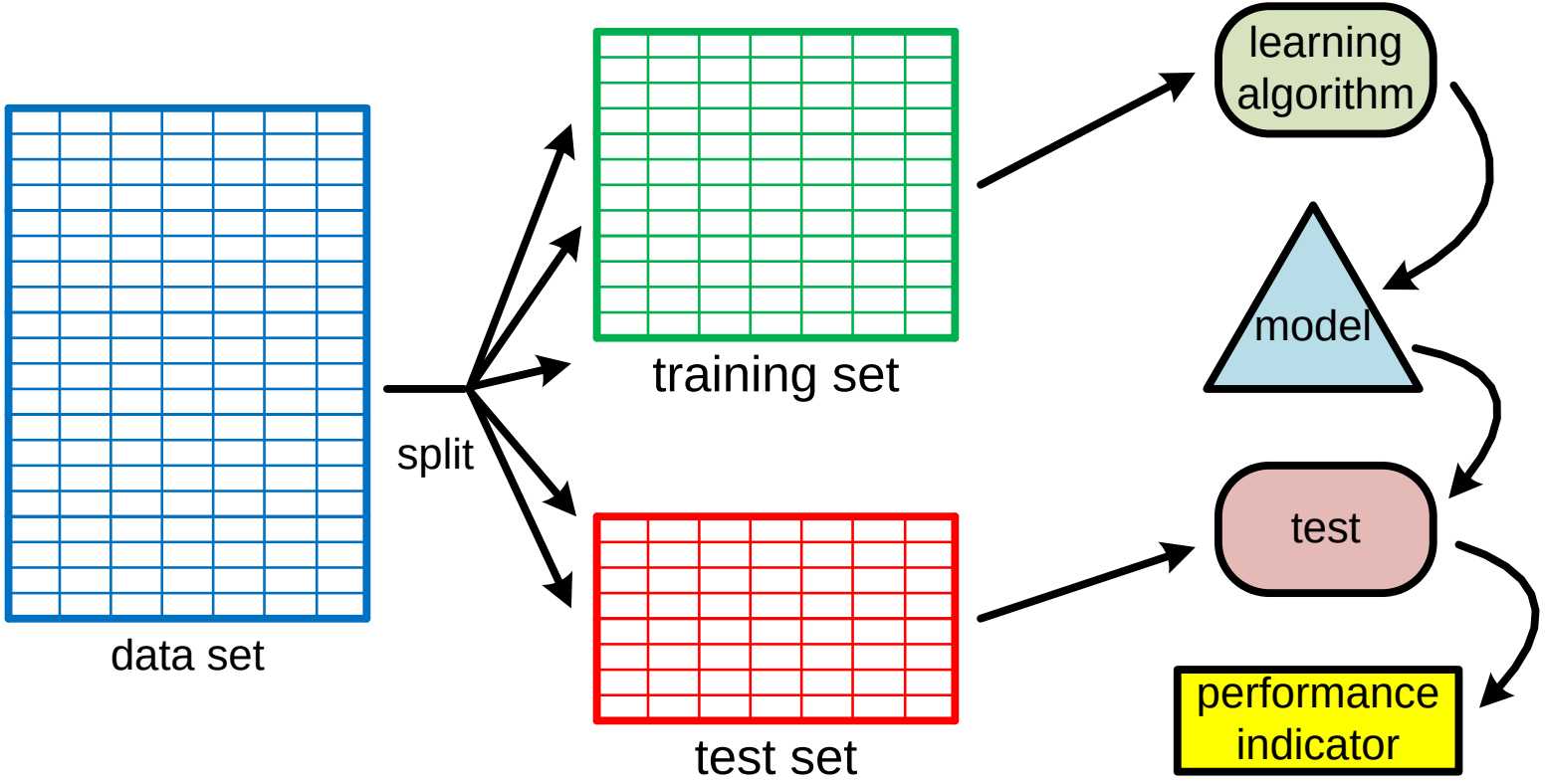
… but the model will be overfitting the data set and will most likely not apply to any new friends.

Overfitting and underfitting

- **Overfitting**: the model is too specific for the data set used to learn the model and performs poorly on new instances.
  - If birth date = 16-05-1998 and eye color = blue, then length = 172.8 cm.
- **Underfitting**: the model is too general and does not exploit the data.
  - If gender = male, then length > 1 meter.

©Wil van der Aalst & TU/e (use only with permission & acknowledgements)

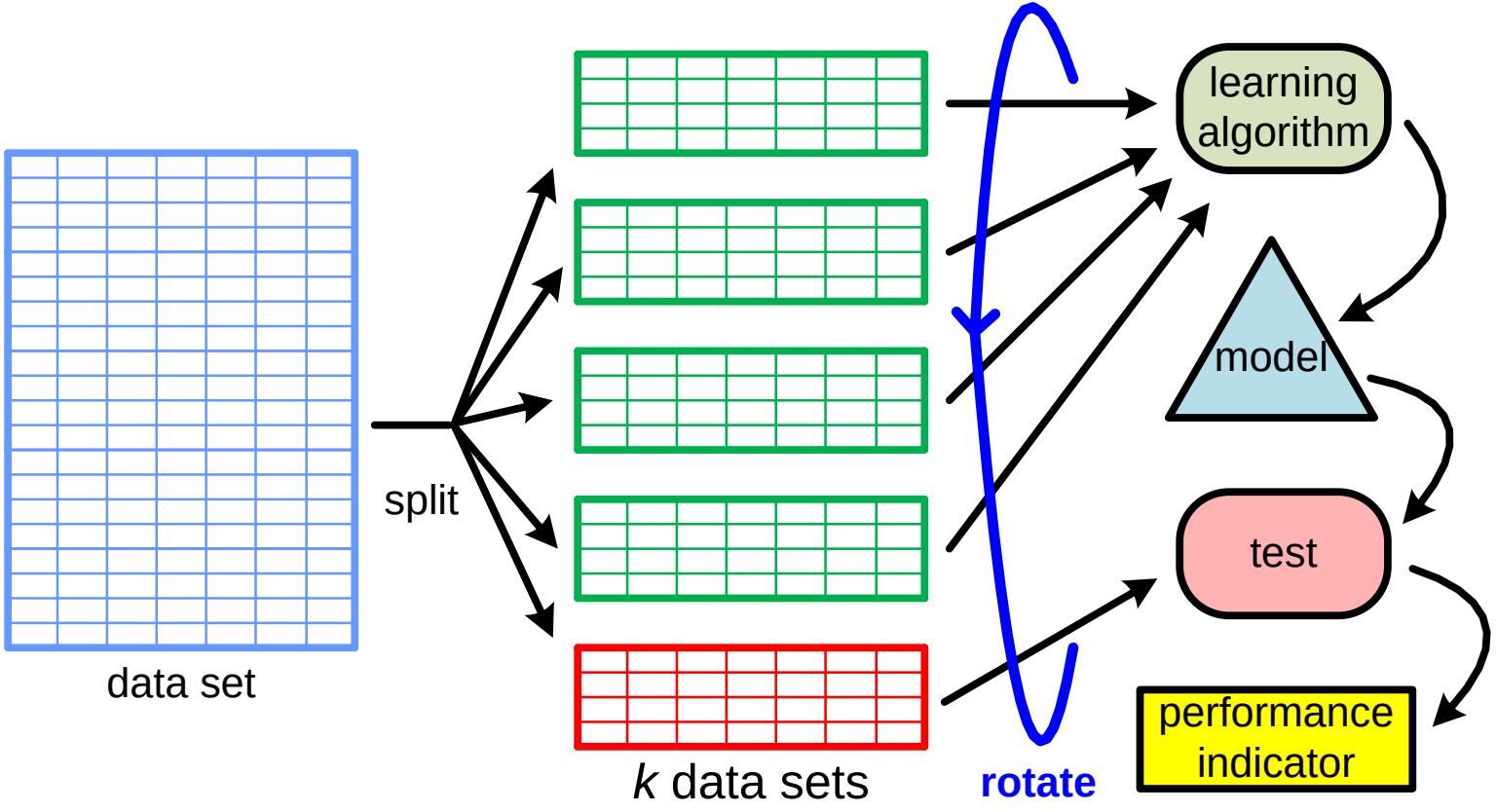
Cross-validation



©Wil van der Aalst & TU/e (use only with permission & acknowledgements)

TU/e

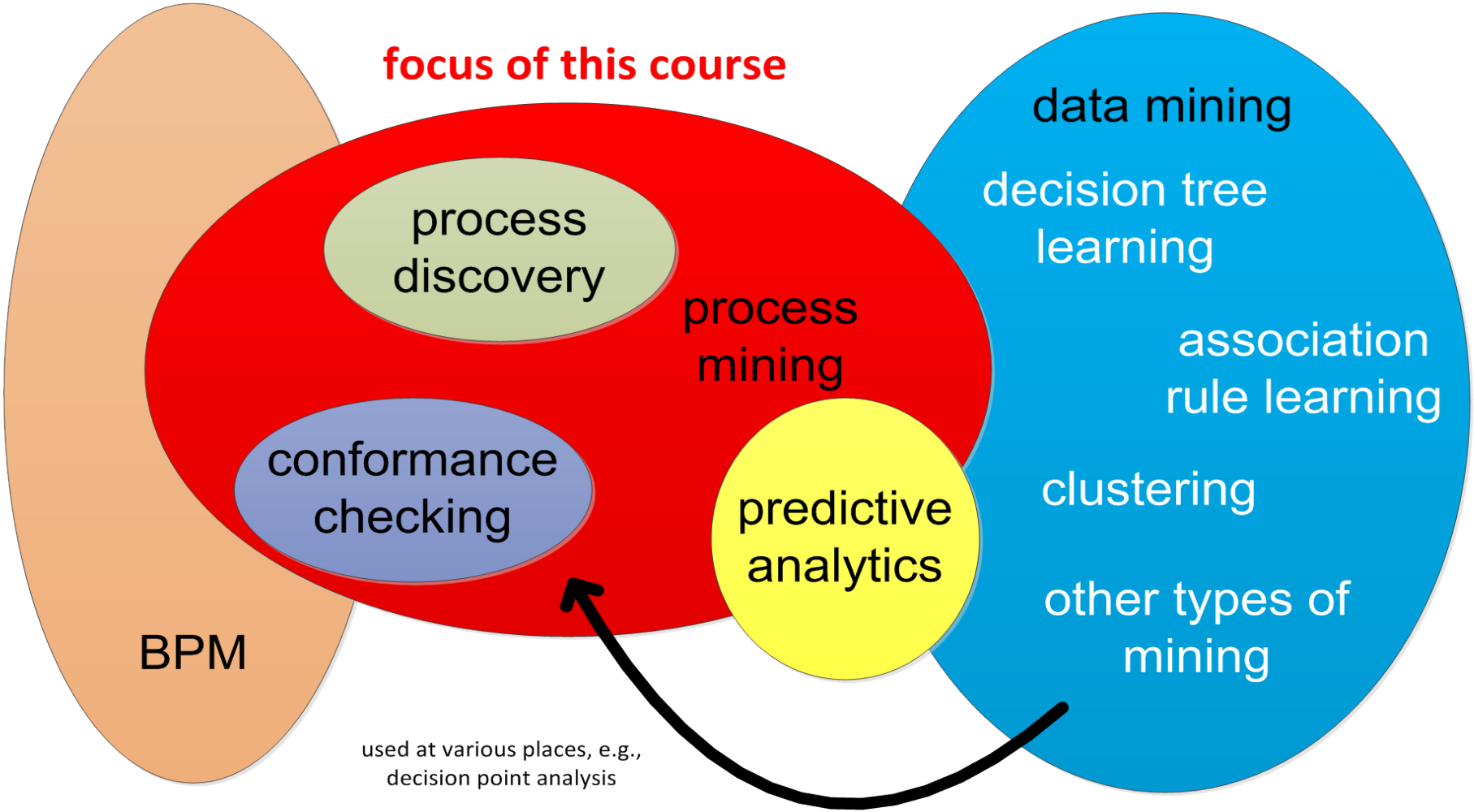
k-fold cross-validation

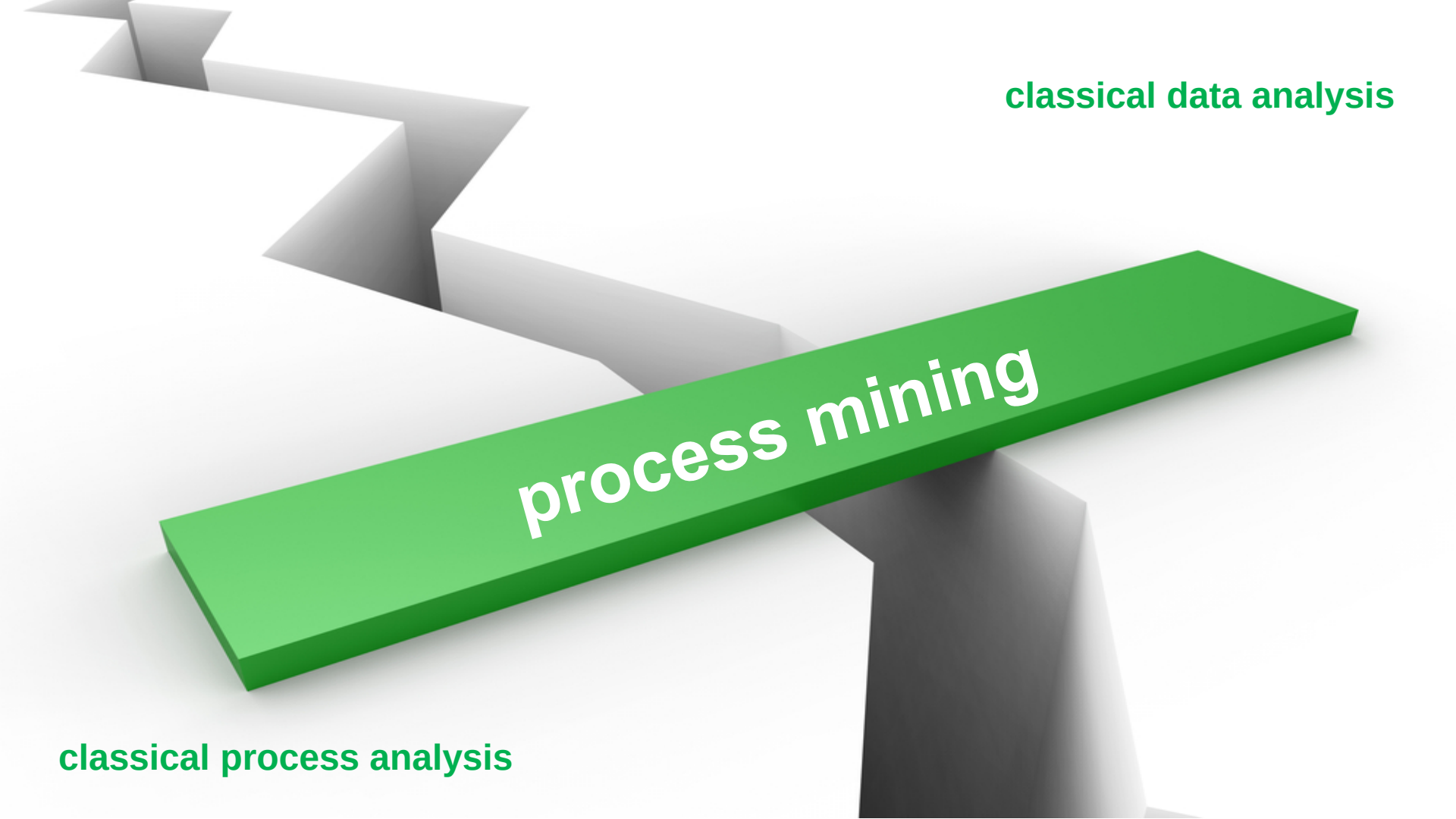


©Wil van der Aalst & TU/e (use only with permission & acknowledgements)

Possible complications

- Concept drift (model should change over time).
- No negative examples (we only know about sick customers that complained afterwards).
- …


focus of this course
data mining
decision tree learning
association rule learning
clustering
other types of mining
process discovery
process mining
conformance checking
predictive analytics
BPM
used at various places, e.g., decision point analysis


classical data analysis
process mining
classical process analysis

Part I: Preliminaries

- Chapter 1: Introduction
- Chapter 2: Process Modeling and Analysis
- Chapter 3: Data Mining

Part II: From Event Logs to Process Models

- Chapter 4: Getting the Data
- Chapter 5: Process Discovery: An Introduction
- Chapter 6: Advanced Process Discovery Techniques

Part III: Beyond Process Discovery

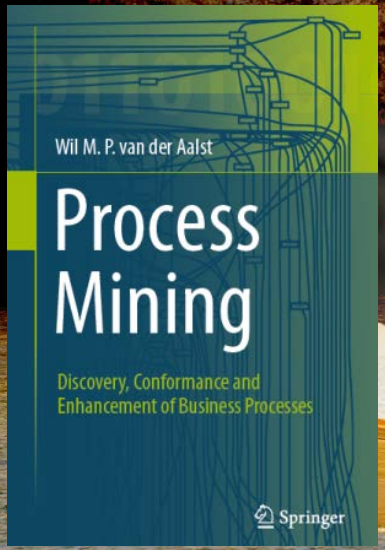
- Chapter 7: Conformance Checking
- Chapter 8: Mining Additional Perspectives
- Chapter 9: Operational Support

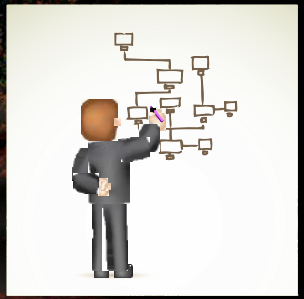
Part IV: Putting Process Mining to Work

- Chapter 10: Tool Support
- Chapter 11: Analyzing "Lasagna Processes"
- Chapter 12: Analyzing "Spaghetti Processes"

Part V: Reflection

- Chapter 13: Cartography and Navigation
- Chapter 14: Epilogue





©Wil van der Aalst & TU/e (use only with permission & acknowledgements)

TU/e