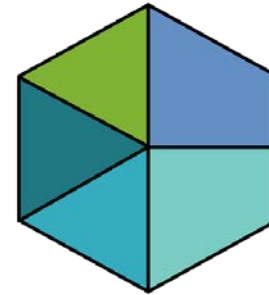




**Hochschule
Kaiserslautern**
University of
Applied Sciences

Angewandte
Ingenieurwissenschaften
Kaiserslautern



elab2go
Mobile
Engineering
Lab

Predictive Maintenance: Wie wird eine Datenanalyse im Rahmen der vorausschauenden Wartung durchgeführt?


Prof. Dr. Eva Maria Kiss

M. Sc. Anke Welz

25.07.2019, Kaiserslautern

Wer sind wir?

elab2go - Mobile Engineering Lab

Die **elab2go-Plattform**  <https://www.elab2go.de>
veranschaulicht neue Technologien und Trends
anhand didaktisch aufbereiteter Demonstratoren.

Wir sind Partner im Innovationsbereich Produkte der **Offenen Digitalisierungsallianz Pfalz** und in das **Engineering 4.0 Lab** Kaiserslautern [e4lab](https://www.e4lab.de) eingebunden.



Wir vermitteln einen Einstieg in praktische Anwendungen im Kontext Industrie 4.0 (Internet der Dinge, Predictive Maintenance) mittels Live-Vorstellung und interaktiven Apps.



Was steckt hinter dem Begriff Predictive Maintenance?

Der Begriff Predictive Maintenance kann mit „Vorausschauender Wartung“ übersetzt werden und ist eine praktische Anwendung im Kontext Industrie 4.0 und ein greifbares Beispiel für das „Internet der Dinge“ (Internet of Things, IoT).

Wieso ein Beispiel für das IoT?

- Digitale Klone von Anlagen/Maschinen erstellen
- Konkret: Erhebung und Speicherung von Mess- und Produktionsdaten → Information über den aktuellen Zustand
- Zustand in Ordnung oder Empfehlung der Wartung.
→ „Vorausschauende Wartung“.



Welche Vorteile bietet die Vorausschauende Wartung?

Durch die Vorhersage der Wartung

- lassen sich ungeplante Maschinenausfälle vermeiden und
- Ressourcen für Instandhaltungsarbeiten besser planen, dies **verringert Kosten** und **verbessert die Produktivität** der Maschinen.

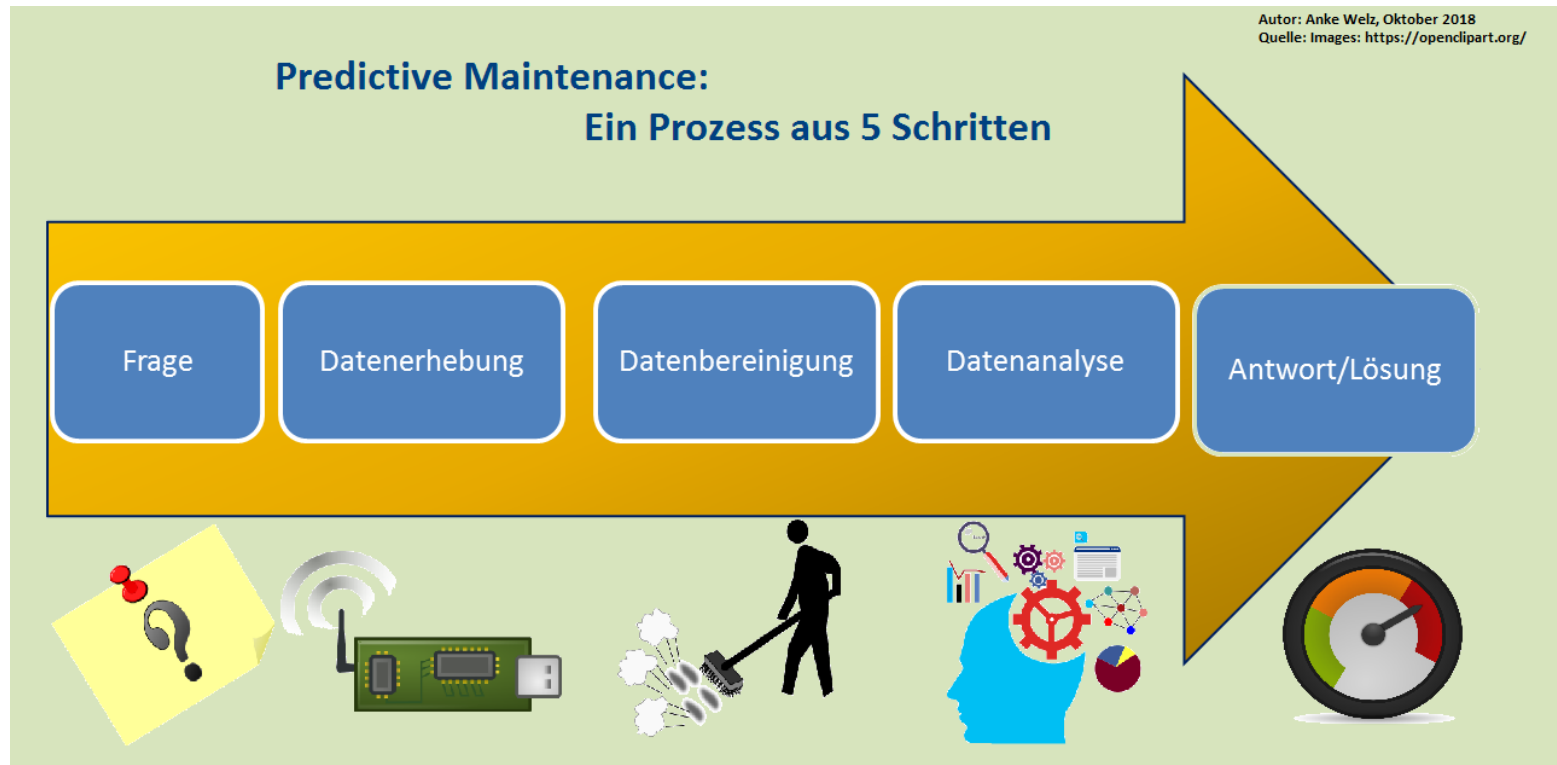
Durch die andauernde Überwachung mittels der bereits installierten Sensoren

- kann das Wartungssystem verbessert und
- immer besser an die zu überwachende Anlage angepasst werden.



Predictive Maintenance in fünf Schritten

Predictive Maintenance wird in vielen Unternehmen als ein Prozess in fünf Schritten abgebildet.



Predictive Maintenance in fünf Schritten

Schritt 1: Die Fragestellung

Suche eines Verfahrens, das frühzeitig vor dem Ausfall einer Maschine warnt und somit Ausfallzeiten und Kosten niedrig hält.

Soll das Gesamt- oder nur Teilsystem überwacht werden?

Wann (bei welchen Messwerten) sollte ein Bauteil ausgewechselt werden?

Schritt 2: Datenerhebung

Die Anlage wird mit Sensoren ausgestattet, die die nötigen Werte (Temperatur, Laufbewegung, Druck, usw.) an den entsprechenden Bauteilen erfassen und speichern.



Predictive Maintenance in fünf Schritten

Schritt 3: Datenbereinigung

Erfahrungsgemäß müssen die gesammelten Daten nun von denen bereinigt werden, die Fehlmessungen oder nicht vollständige Messungen darstellen. Dies erfolgt als Vorarbeit der Datenanalyse

Schritt 4: Datenanalyse

Auf die bereinigten Daten wird nun die Datenanalyse mit Hilfe von Algorithmen und Verfahren des maschinellen Lernens angewandt, um die Fragestellung aus Schritt 1 zu beantworten. Mehr dazu später.



Predictive Maintenance in fünf Schritten

Schritt 5: Die Antwort/Lösung

Die Ausgabe der Analyse-Software wird nun von dem Analysten in eine verständliche Antwort übersetzt und ein für das Unternehmen **leicht anwendbares Wartungssystem** entwickelt. Es kann z.B. ein **Ampelsystem** entstehen, das je nach Zustand des Bauteils die Farbe grün/gelb/rot anzeigt:

Grün steht für die Funktionsfähigkeit, Gelb für die Beobachtung und Rot für den empfohlenen Austausch des Bauteils



Datenanalyse im Rahmen der Pred. Maint.

- Anwendung von Algorithmen und Verfahren des maschinellen Lernens
- Dabei können statistische Modelle oder auch Entscheidungsbäume zur Lösung führen.
- Mit Hilfe von Testverfahren wie der Kreuzvalidierung werden Kennzahlen (Vertrauenswahrscheinlichkeit, Genauigkeit) bestimmt, mit deren Hilfe man die Güte des Vorhersagemodells bewerten kann.
- Auf Basis von Schwellwerten die Eintrittswahrscheinlichkeit für den Ausfall eines Bauteils beantwortet.



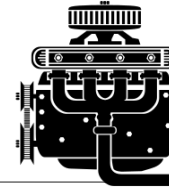
Datenanalyse im Rahmen der Pred. Maint.

Die genauen Schritte zur Beantwortung einer Frage im Rahmen von Predictive Maintenance mittels Datenerhebung und -verarbeitung, Analyse und die Formulierung einer verständlichen Antwort werden nun durch **verschiedene Demonstratoren anhand eines Beispiels aus der Automobilbranche** erklärt.

Die Demonstratoren sind **interaktiv** und die **Anpassungsmöglichkeiten** innerhalb der Demos veranschaulichen direkt die Auswirkungen auf die Antwort Wartung JA/NEIN.



Der Automobildatensatz



An 7 Motoren wurden mittels Sensoren

- Temperatur- und Druckmessungen
- sowie Mengenangaben zum Kraftstoff und zu Abgasdämpfen erfasst und zusätzlich auch, ob der **Motor ausgefallen ist oder nicht**. Insgesamt: 22 Merkmale in jeder der 136 Beobachtungen

Messungsnummer	Ausfall	Kuehlmitteltemperatur	Einspritzmenge	Kurzzeit	Luftdruck	Katalysatortemp	Kategorie
M_001	nein	2,6		,9	2,8		normal
M_002	ja	9,2		22,7	2,8		hoch
M_003	nein	3,2		28,5	,1		normal
M_004	ja	2,8		2,6	,2		normal
M_005	ja	4,9		14,4	3,0		hoch
M_006	nein	,8		5,5	2,9		hoch
M_007	nein	6,2		26,4	,1		normal
		⋮					
M_133	nein	2,0		3,6	,2		normal
M_134	nein	1,5		1,2	1,4		normal
M_135	nein	4,4		5,3	2,6		hoch
M_136	nein	3,1		5,4	2,6		hoch



Datenanalyse - Die Fragestellung

Welche Kombination von Merkmalen wird zu einem Ausfall führen?

Messungsnummer	Ausfall	Kuehlmitteltemperatur	Einspritzmenge	Kurzzeit	Luftdruck	Katalysatortemp	Kategorie
M_005	ja	4,9		14,4	3,0		hoch
M_006	nein	,8		5,5	2,9		hoch
M_007	nein	6,2		26,4	,1		normal
		⋮					
M_133	nein	2,0		3,6	,2		normal
M_134	nein	1,5		1,2	1,4		normal

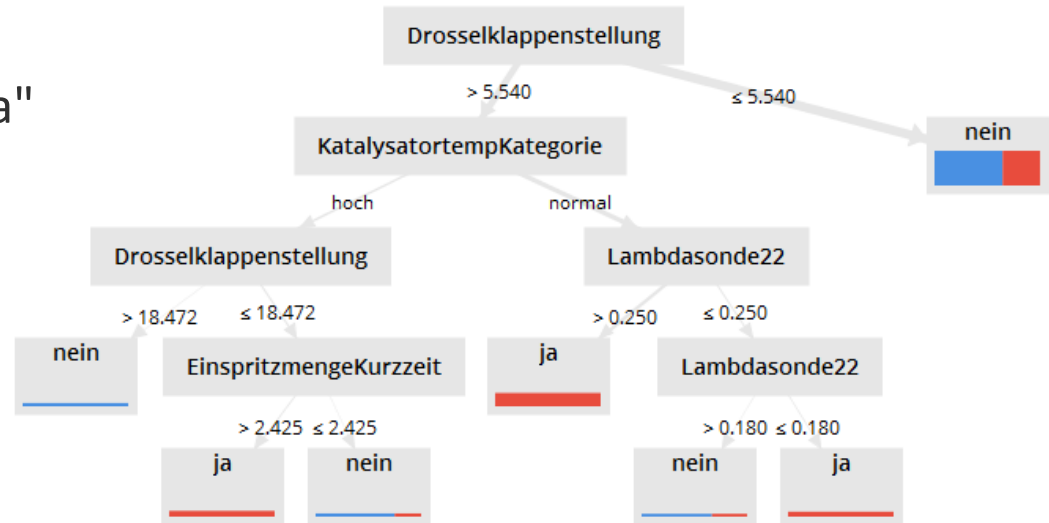
**Ein Entscheidungsbaum
als Vorhersagemodell
liefert eine übersichtliche
und schnell interpretierbare
Antwort ...**



Datenanalyse – Der Entscheidungsbaum

Welche Kombination von Merkmalen wird zu einem Ausfall führen?

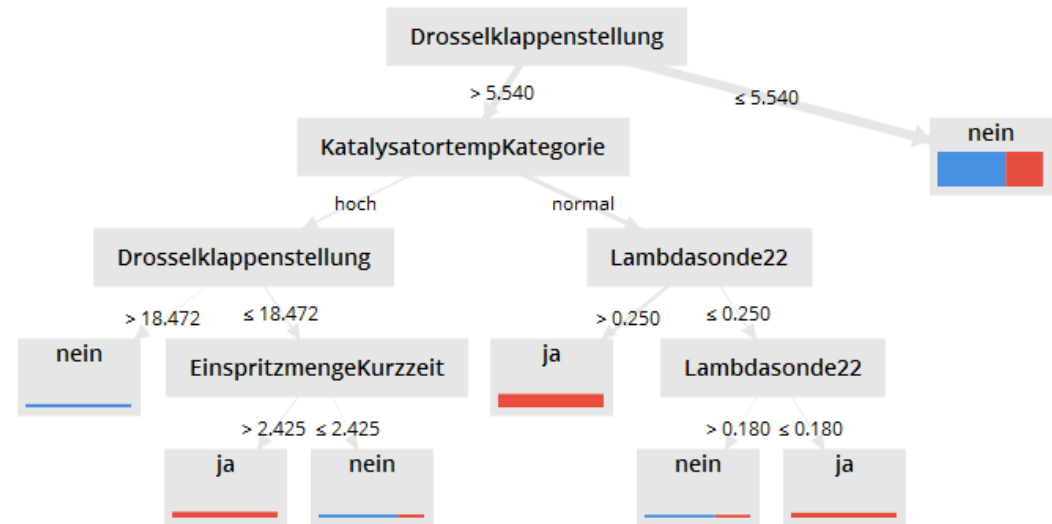
Wenn am Ende eines Pfades "Ja" steht, dann führt der Pfad/ die Merkmalskombination zu einem Ausfall des Motors andernfalls, d.h. bei "Nein", kommt es zu keinem Ausfall.



Datenanalyse – Die Vorhersage

Welche Kombination von Merkmalen wird zu einem Ausfall führen?

Eine Vorhersage für neue Merkmals-kombinationen ist nun möglich: **Durchlaufe den passenden Pfad.**

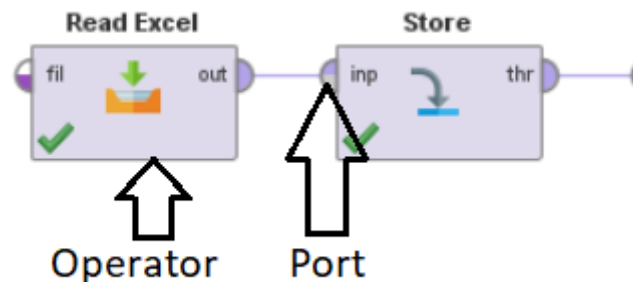


Wir verwenden **verschiedene Tools (RapidMiner, RStudio, MATLAB)** um einen Entscheidungsbaum als Vorhersagemodell zu erstellen, mit dessen Hilfe dann Vorhersagen über den Ausfall eines Motors getroffen werden können.



Datenanalyse mit RapidMiner – Die Entwicklungsumgebung

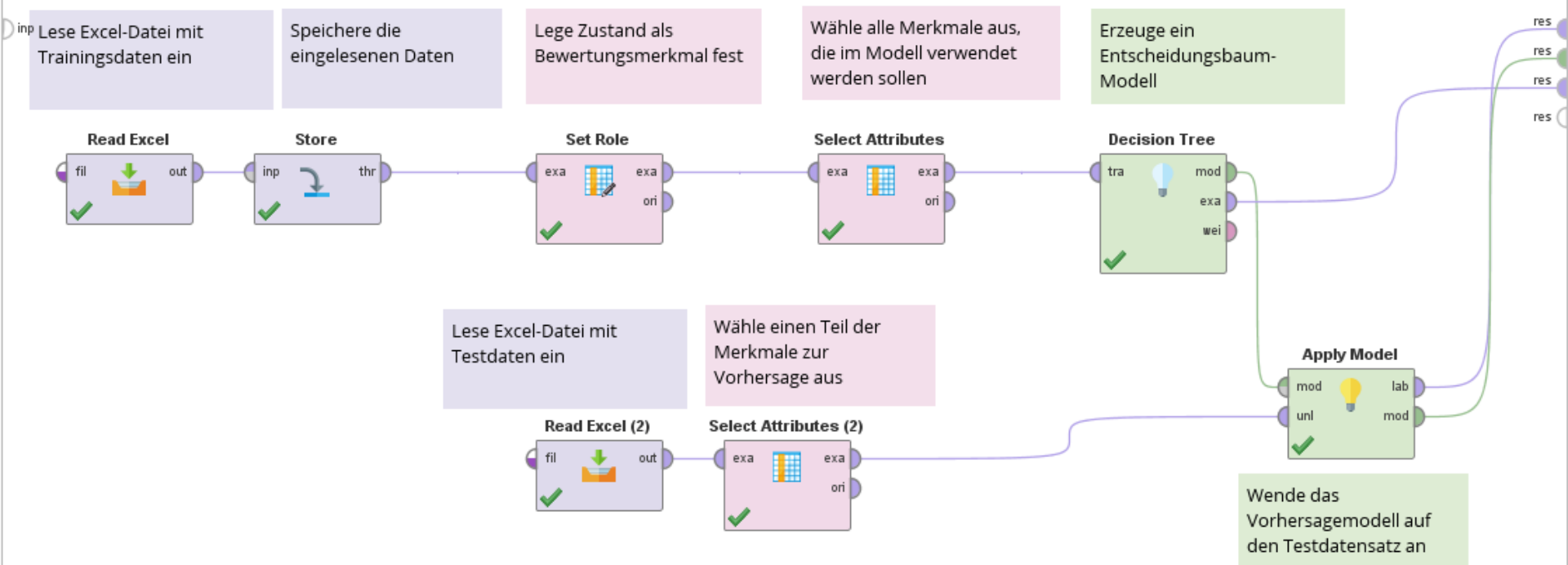
- RapidMiner ist eine Entwicklungsumgebung für Datenanalyse, die als Open Source-Software an der Uni Dortmund entstanden ist und heute zu den Marktführern gehört.
- RapidMiner bildet den Datenanalyse-Prozess mit Hilfe von Operatoren ab, die alle gängigen Aufgaben der Datenanalyse abbilden und über Ports miteinander verknüpft werden.



Datenanalyse mit RapidMiner – Der Datenanalyse-Prozess

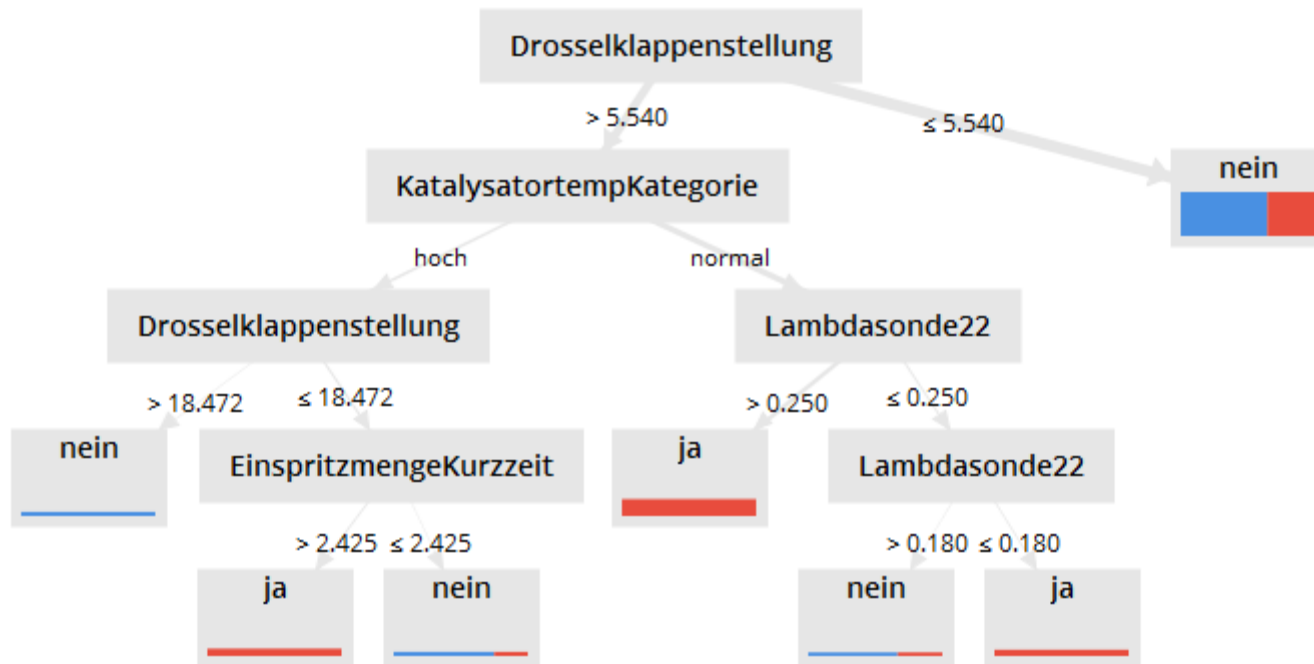
Der Prozessverlauf beinhaltet verschiedene Schritte:

- Einlesen und Speichern der Daten (lila-markiert),
- Bearbeitung der Daten (rosa-markiert),
- die Analyse der Trainingsdaten und die Vorhersage mit Hilfe des erstellten Entscheidungsbaum-Modells (grün-markiert).



Datenanalyse mit RapidMiner – Der Entscheidungsbaum

Der Entscheidungsbaum wählt nur einen Teil der 22 Merkmale zur Vorhersage eines Ausfalls aus.



Datenanalyse mit RapidMiner – Der Entscheidungsbaum

- Je höher das Merkmal im Baum steht, d.h. je weiter oben dessen Position ist, desto mehr Einfluss hat es auf die Vorhersage eines Ausfalls.
- In unserem Baum hat die **Drosselklappenstellung den höchsten Einfluss**, da bei einem Messwert kleiner oder gleich 5.54 ein Ausfall direkt ausgeschlossen ist.
- Eine alternative Darstellung des Entscheidungsbaum-Modells ist als Textform gegeben, einige Werkzeuge zur Predictive Maintenance geben diese Darstellung standardmäßig aus.



Datenanalyse mit RapidMiner – Der Entscheidungsbaum

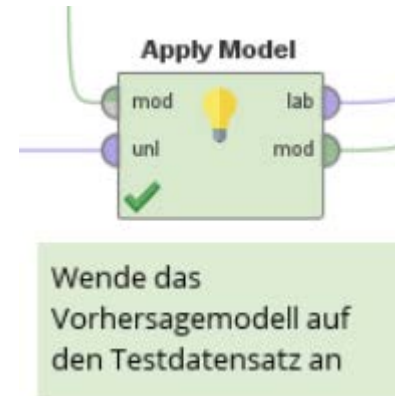
Tree

```
Drosselklappenstellung > 5.540
|   KatalysatortempKategorie = hoch
|   |   Drosselklappenstellung > 18.472: nein {nein=4, ja=0}
|   |   Drosselklappenstellung ≤ 18.472
|   |   |   EinspritzmengeKurzzeit > 2.425: ja {nein=0, ja=10}
|   |   |   EinspritzmengeKurzzeit ≤ 2.425: nein {nein=3, ja=1}
|   KatalysatortempKategorie = normal
|   |   Lambdasonde22 > 0.250: ja {nein=0, ja=28}
|   |   Lambdasonde22 ≤ 0.250
|   |   |   Lambdasonde22 > 0.180: nein {nein=2, ja=1}
|   |   |   Lambdasonde22 ≤ 0.180: ja {nein=0, ja=8}
Drosselklappenstellung ≤ 5.540: nein {nein=51, ja=28}
```



Datenanalyse mit RapidMiner – Die Vorhersage

Die Vorhersage für neue Merkmalskombinationen erfolgt mittels des „Apply Model“-Operators:



Die Vorhersage (prediction) wird in einer Tabelle ausgegeben:

ExampleSet (4 examples, 3 special attributes, 23 regular attributes)

Filter (4 / 4 examples):

Row No.	prediction(Ausfall)	confidence(nein)	confidence(ja)	Ausfall	Einspritzme...	Einspritzme...	Kraftstoffdr...
1	nein	0.667	0.333	?	0.918	4.229	13.792
2	nein	0.646	0.354	?	22.732	15.307	8.553
3	nein	0.646	0.354	?	28.526	8.735	4.898
4	ja	0	1	?	2.642	5.857	7.417



Datenanalyse - Zwischenergebnis

- Wir haben wir die Frage nach den Kombinationen von Merkmalen, die zu einem Ausfall eines Motors führen, anhand des Vorhersagemodells des Entscheidungsbaumes beantwortet.
- Dieser Baum wurde danach interpretiert und zur Vorhersage eines Ausfalls bei neu erfassten Daten verwendet.

→ Nun stellt sich aber die Frage wie gut das Vorhersagemodell überhaupt ist.



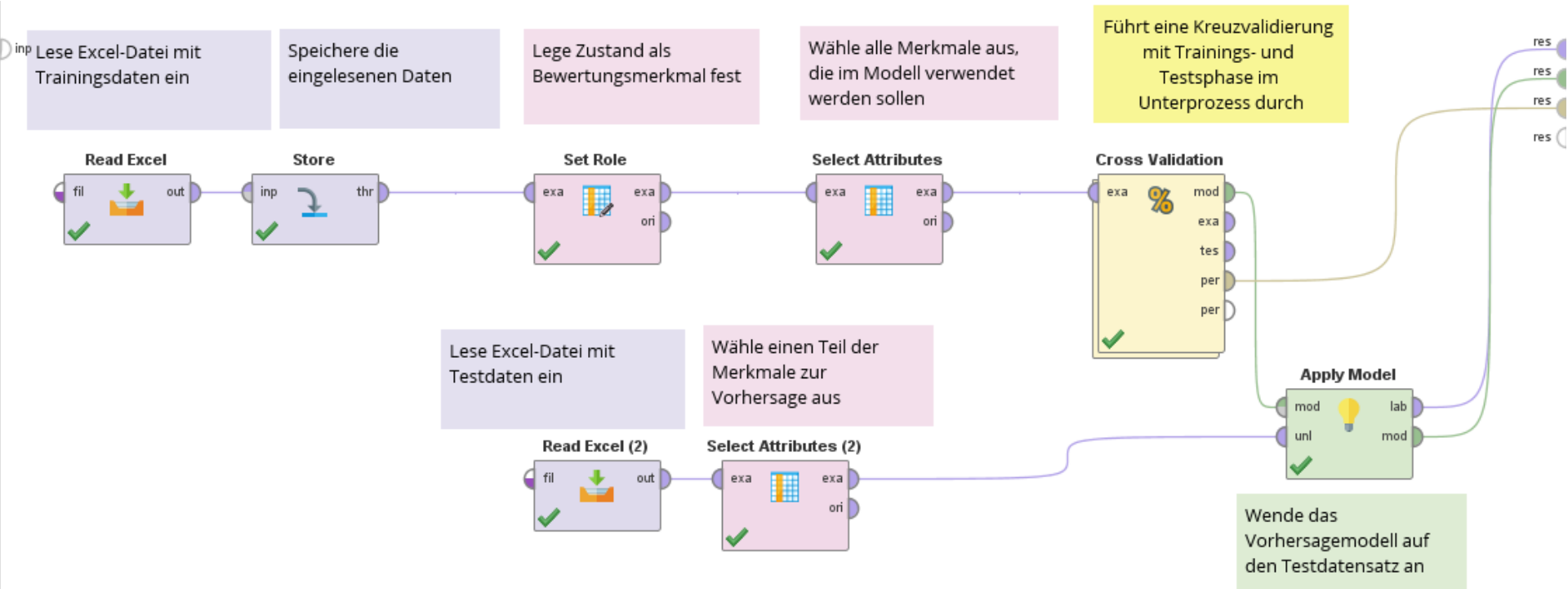
Datenanalyse im Rahmen der Pred. Maint. – Fortsetzung

- Uns interessiert, wie sicher das Vorhersagemodell den Ausfall eines Motors vorhersagt.
- Anhand der Kennzahlen zum Testfehler lässt sich dies beantworten: Liegen die Kennzahlen in einem vorher festgelegten, akzeptablen Bereich? Wenn ja, dann ist das Modell gut.
- Bei einem guten Modell können wir uns auf die Aussagen und Vorhersagen durch das Modell verlassen.



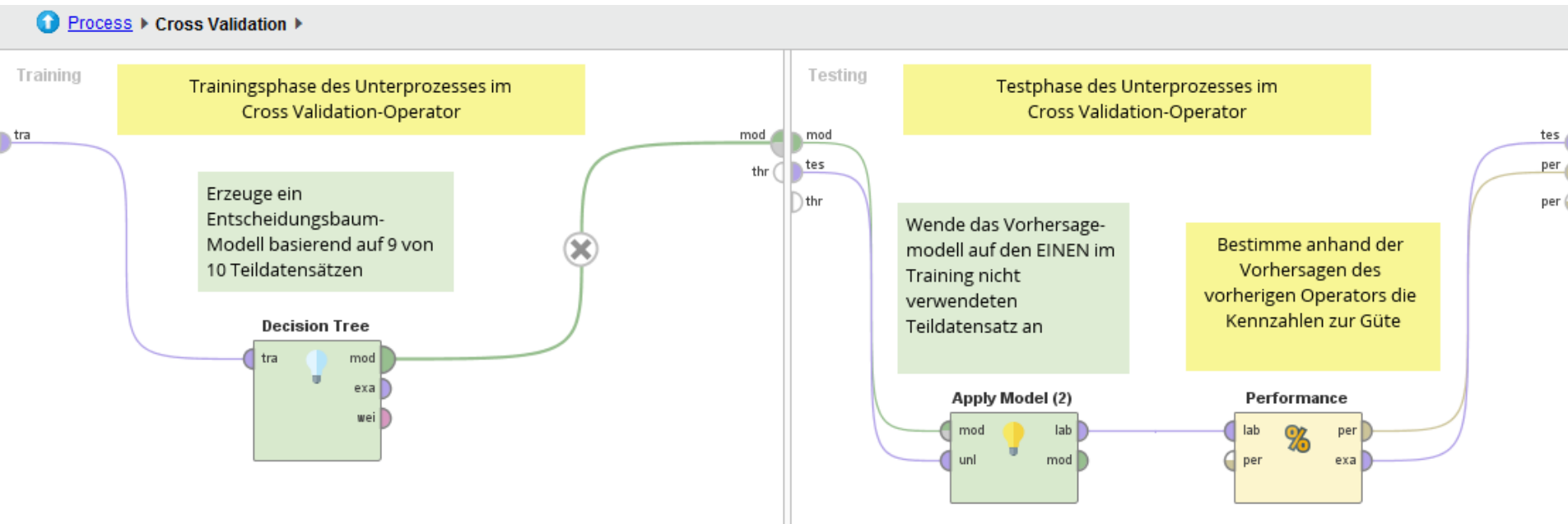
Datenanalyse – Die Kennzahlen mit RapidMiner

Der zur Erstellung des Entscheidungsbaums verwendete RapidMiner-Prozess wird nun erweitert, um zusätzlich die zur Bewertung der Güte des Vorhersagemodells benötigten Kennzahlen zum Testfehler zu bestimmen.



Datenanalyse – Die Kennzahlen mit RapidMiner

Die Schritte "Einlesen und Speichern der Daten" (lila-markiert) und "Bearbeitung der Daten" (rosa-markiert) ändern sich nicht, einzig der Operator "DecisionTree" des Datenanalyse-Prozesses wird auf oberster Ebene des Prozesses durch den "CrossValidation"-Operator (gelb-markiert) ersetzt.



Datenanalyse – Die Beurteilung des Modells

Wie gut der Entscheidungsbaum zur Vorhersage eines Ausfalls geeignet ist, können wir nun anhand der Kennzahlen zum Testfehler beurteilen:

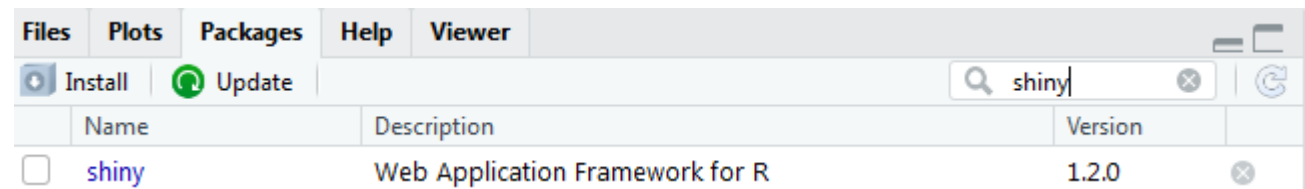
- **Accuracy = 61.54%**
- **Recall (True Positive Rate) TPR = 59.29%**
- **Precision = 71.79%**
- **AUC = 52.3%**

Bei einem „guten“ Modell liegen die Kennzahlen „nahe“ an 100%. → Hier liegen die Kennzahlen in einem moderaten Bereich und können mit Vorbehalt zur Vorhersage eines Ausfalls verwendet werden.



Datenanalyse mit RStudio – Die Entwicklungsumgebung

- RStudio ist eine Entwicklungsumgebung und grafische Benutzeroberfläche für Datenanalyse mit Hilfe der statistischen Programmiersprache R, die in Open-Source- und kommerziellen Editionen erhältlich ist.
- Die zur Datenanalyse benötigten Algorithmen sind in Pakete zusammengefasst, die nach Bedarf geladen werden können.
- Mit Hilfe des Packages R Shiny aus R können z.B. Webseiten für Datenanalysen gebaut werden. → Demo auf elab2go



```
library(shiny)
library(shinydashboard) #for dashboard design app v1.0
library(shinyBS) #for popovers and tooltips
```



Datenanalyse mit RStudio – Der Quellcode

Im Analyse-Code werden die benötigten Pakete geladen, z.B. das "rpart"-Paket, dieses enthält Funktionen, die dann zur Erstellung eines Klassifikationsmodells verwendet werden.

```
1 #Demo 3: Predictive Maintenance mit R
2 #     Analyse
3
4 # Lade benötigte Pakete
5 library(rpart)
6
7 #Setze den Arbeitspfad
8 setwd("C:/Demo3")
9
10 #Lese CSV-Datei und speichere
11 #sie unter dem Namen "demo3"
12 demo3=read.csv2( "demo3_data.csv")
13
14 #Wähle Merkmale aus
15 #Erstelle Trainingsdaten demo3training
16 demo3training=demo3[ ,c(-1,-24)]
17
18 #Erstelle Entscheidungsbaum
19 #mit dem CART-Algorithmus
20 fit = rpart(Ausfall ~ .,
21           data=demo3training,
22           method="class",
23           parms=list( split="information"),
24           control=list( minsplit=15))
25
26 #Gebe den Baum in Textform aus
27 print(fit)
28
29 #Erstelle die Grafik des Baums
30 plot(fit, uniform=TRUE,
31       main="Class. Tree for Automotive")
32 text(fit, use.n=TRUE, all=TRUE, cex=.8, pretty=FALSE)
```

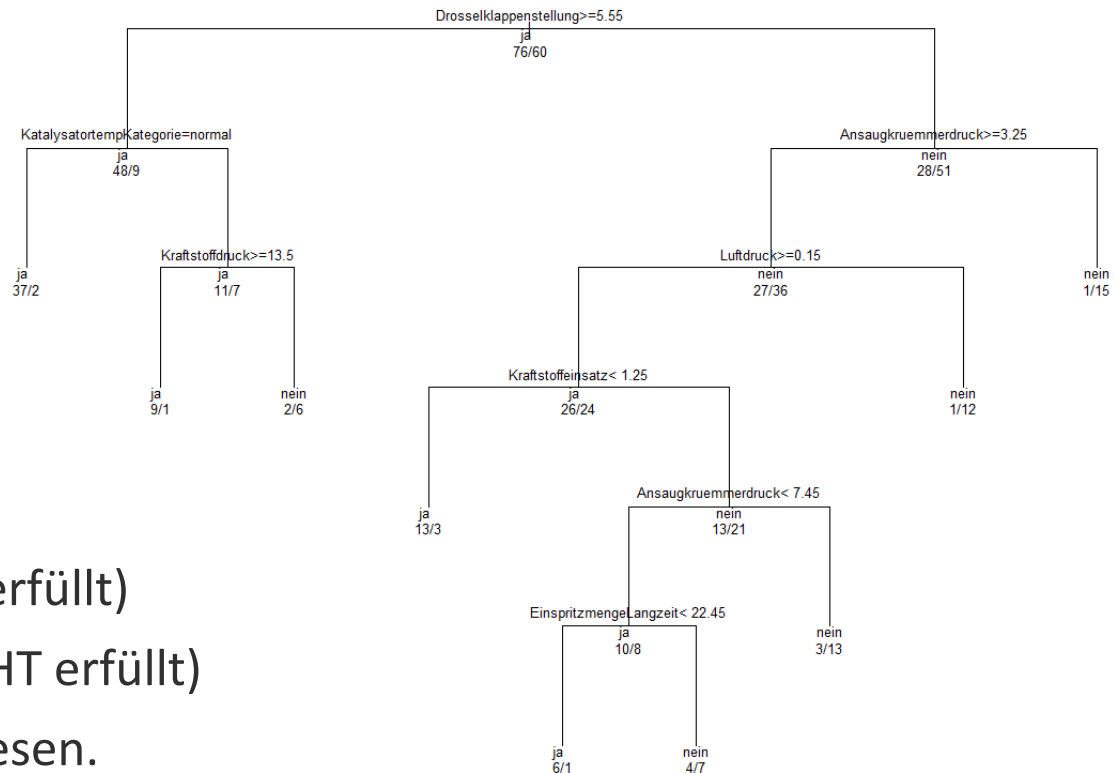


Datenanalyse mit RStudio – Der Entscheidungsbaum

Der Entscheidungsbaum wählt nur einen Teil der 22 Merkmale zur Vorhersage eines Ausfalls aus.

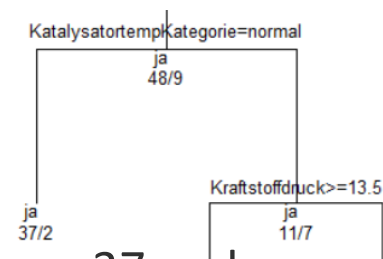
An jedem Knoten wird für das jeweilige Merkmal eine Bedingung abgefragt und je nach Beobachtungswert wird nach links (Bedingung erfüllt) oder rechts (Bedingung NICHT erfüllt) zum weiteren Knoten verwiesen.

Entscheidungsbaum für den Automobildatensatz



Datenanalyse mit RStudio – Der Entscheidungsbaum

- Je höher das Merkmal im Baum steht, d.h. je weiter oben dessen Position ist, desto mehr Einfluss hat es auf die Vorhersage eines Ausfalls, in unserem Baum hat die **Drosselklappenstellung den höchsten Einfluss**.
- An den Aufteilungen in Ja/Nein in den Blättern erkennt man, wie viele Beobachtungen des Trainingsdatensatzes mit beobachteten Ausfällen Ja/Nein diesem Blatt zugeordnet werden.



- Z.B. insgesamt 39 Beobachtungen im ganzen linken Ast, bei denen 37 mal kein Ausfall vorliegt und 2 mal eine Ausfall. Daraus lassen sich die **Wahrscheinlichkeiten für Ausfall Ja oder Nein** je Blatt bestimmen, die auch bei der Ausgabe des Baums **in Textform** angegeben sind.



Datenanalyse mit RStudio – Der Entscheidungsbaum

```
> print(fit) #Gebe den Baum in Textform aus
n= 136
```

```
node), split, n, loss, yval, (yprob)
* denotes terminal node
```

```
1) root 136 60 ja (0.55882353 0.44117647)
  2) Drosselklappenstellung>=5.55 57 9 ja (0.84210526 0.15789474)
    4) KatalysatortempKategorie=normal 39 2 ja (0.94871795 0.05128205) *
    5) KatalysatortempKategorie=hoch 18 7 ja (0.61111111 0.38888889)
      10) Kraftstoffdruck>=13.5 10 1 ja (0.90000000 0.10000000) *
      11) Kraftstoffdruck< 13.5 8 2 nein (0.25000000 0.75000000) *
  3) Drosselklappenstellung< 5.55 79 28 nein (0.35443038 0.64556962)
    6) Ansaugkruemmerdruck>=3.25 63 27 nein (0.42857143 0.57142857)
      12) Luftdruck>=0.15 50 24 ja (0.52000000 0.48000000)
        24) Kraftstoffeinsatz< 1.25 16 3 ja (0.81250000 0.18750000) *
        25) Kraftstoffeinsatz>=1.25 34 13 nein (0.38235294 0.61764706)
          50) Ansaugkruemmerdruck< 7.45 18 8 ja (0.55555556 0.44444444)
            100) EinspritzmengeLangzeit< 22.45 7 1 ja (0.85714286 0.14285714) *
            101) EinspritzmengeLangzeit>=22.45 11 4 nein (0.36363636 0.63636364) *
          51) Ansaugkruemmerdruck>=7.45 16 3 nein (0.18750000 0.81250000) *
      13) Luftdruck< 0.15 13 1 nein (0.07692308 0.92307692) *
    7) Ansaugkruemmerdruck< 3.25 16 1 nein (0.06250000 0.93750000) *
```



Datenanalyse mit RStudio – Die Vorhersage

Die Vorhersage für neue Merkmalskombinationen erfolgt mittels der „predict“-Funktion:

```
#Vorhersage für die Testdaten/neue Beobachtungen
pred = predict(fit, newdata=testdata, type="class")
```

Die Vorhersage wird als Vektor/Matrix ausgegeben:

```
> #Vorhersage für die Testdaten/neue Beobachtungen
> predict(fit, newdata=testdata, type="class")
  1    2    3    4
ja   ja nein  ja
Levels: ja nein
>
> #erweiterte Vorhersage für die Testdaten/neue Beobachtungen
> cbind(predict(fit, newdata=testdata, type="prob"),predict(fit, newdata=testdata, type="class"))
      ja      nein  1  2
1 0.94871795 0.05128205 1  1
2 0.81250000 0.18750000 1  1
3 0.07692308 0.92307692 2  2
4 0.94871795 0.05128205 1  1
```

1=ja
2=nein



Datenanalyse – Zwischenergebnis

- Wir haben wir die Frage nach den Kombinationen von Merkmalen, die zu einem Ausfall eines Motors führen, anhand des Vorhersagemodells des Entscheidungsbaumes beantwortet.
- Dieser Baum wurde danach interpretiert und zur Vorhersage eines Ausfalls bei neu erfassten Daten verwendet.

→ Nun stellt sich aber die Frage wie gut das Vorhersagemodell überhaupt ist.



Datenanalyse im Rahmen der Pred. Maint. – Fortsetzung

- Uns interessiert, wie sicher das Vorhersagemodell den Ausfall eines Motors vorhersagt.
- Anhand der Kennzahlen zum Testfehler lässt sich dies beantworten: Liegen die Kennzahlen in einem vorher festgelegten, akzeptablen Bereich? Wenn ja, dann ist das Modell gut.
- Bei einem guten Modell können wir uns auf die Aussagen und Vorhersagen durch das Modell verlassen.



Datenanalyse – Die Kennzahlen mit RStudio

- Das "caret"-Paket stellt die Funktion "confusionMatrix" bereit, die fast alle Kennzahlen bestimmt, mit der "performance"-Funktion des "ROCR"-Paketes wird eine weitere Kennzahl zum Testfehler zur Bewertung der Güte des Vorhersagemodells bestimmt.
 - Der Kennzahl-Code besteht aus einer for-Schleife, die bei jedem Schleifen-Durchlauf einen Teildatensatz auswählt, mit den übrigen Teildatensätzen das Vorhersagemodell/Baum erstellt und dann für ausgewählten Teildatensatz eine Vorhersage durchführt.
 - Die vorhergesagten Werte werden dann mit den tatsächlichen Werten für die Variable Ausfall verglichen und somit die Kennzahlen bestimmt.
 - Dies erfolgt K mal und nachdem die Schleife durchlaufen wurde, werden alle K Werte der Kennzahlen gemittelt.
- dies entspricht der Idee der K-fachen Kreuzvalidierung/CrossValidation...



Datenanalyse – Die Kennzahlen mit RStudio



Datenanalyse – Die Beurteilung des Modells

Wie gut der Entscheidungsbaum zur Vorhersage eines Ausfalls geeignet ist, können wir nun anhand der vier ermittelten Kennzahlen zum Testfehler beurteilen:

- **Accuracy = 66.15%**
- **Recall (True Positive Rate) TPR = 69.11%**
- **Precision = 68.74%**
- **AUC = 72.86%**

Bei einem „guten“ Modell liegen die Kennzahlen „nahe“ an 100%. → Hier liegen die Kennzahlen in einem moderaten Bereich und können mit Vorbehalt zur Vorhersage eines Ausfalls verwendet werden.



Predictive Maintenance mit R – Interaktive PredMaintApp

- Ein weiterer Demonstrator des elab2go veranschaulicht die vorherige Datenanalyse anhand einer interaktiven Webanwendung, die mit Hilfe des R-Paketes Shiny entstanden ist .

Welcher Motor wird ausfallen und eine Wartung benötigen?

Die Antwort liefert ein Vorhersage-Modell, dessen Erstellung ein Schritt des Predictive Maintenance-Prozesses ist:

- Auswahl der Merkmale
- Einstellen der Parameter
- Anwendung des Vorhersage-Modells

Auswahl der Merkmale

Im Folgenden können die Merkmale des **Automobildatensatzes** gewählt werden, die zur Vorhersage eines Ausfalls eines Motors im Modell berücksichtigt werden sollen:

Merkmale


Zur Auswahl anklicken:

- Kühlmitteltemperatur
- Einspritzmenge Kurzzeit
- Einspritzmenge Langzeit
- Kraftstoffdruck
- Ansaugkrümmerdruck
- Drosselklappenstellung
- Einlasslufttemperatur
- Lambdasonde11
- Lambdasonde21
- Lambdasonde31
- Lambdasonde41
- Lambdasonde12
- Lambdasonde22
- Lambdasonde32
- Lambdasonde42
- Kraftstoffleitungsdruck
- Abgasrückführung
- Kraftstoffdampfsäuberung
- Kraftstoffeinsatz
- EVAPDruck
- Luftdruck
- Katalysatortemperatur
- Kategorie alle

Parameter

Min. Anz. Beobachtungen pro Knoten:

5 15 50




Mit diesem Parameter wird festgelegt, dass nur Knoten erlaubt sind, zu denen mindestens

15

Maximale Tiefe des Baumes:

1 15



Mit diesem Parameter wird die maximale Tiefe des Baums festgelegt, also wie viele Verzweigungen der Baum maximal enthalten soll.

Merkmalskombinationen führen.

- Durch Eingabebereiche können die Daten und verschiedene Parameter vom Benutzer angepasst werden.

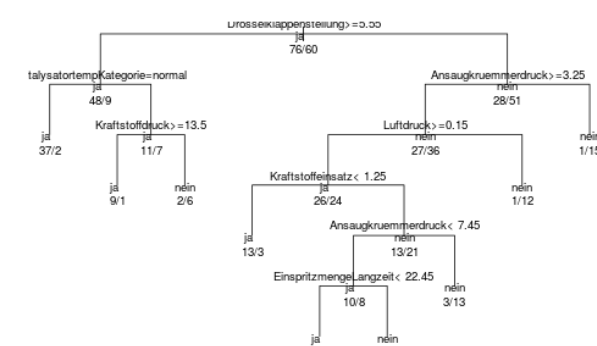


Predictive Maintenance mit R – Interaktive PredMaintApp

Der erstellte Baum

Baum
Textform

Entscheidungsbaum fuer die Automobildaten



Die gewählten Daten

Zeige 5 Beobachtungen Suche:

	Ausfall	Drosselklappenstellung	Kraftstoffeinsatz
1	nein	13.8	1.1
2	ja	3	0.7
3	nein	3	0.9
4	ja	16.2	0.8
5	ja	3.8	2.2

Zeige 1 bis 5 von 136 Beobachtungen
 Vorherige 1 2 3 4 5 ... 28 Naechste

- Ausgabebereiche zeigen den erstellten Entscheidungsbaum, die gewählten Daten und die Vorhersage an.

Die Vorhersage des Ausfalls (JA/NEIN) für die neuen Beobachtungen lautet:

BeobachtungsNr.	Ausfall?
[1,] "1"	"ja"
[2,] "2"	"ja"
[3,] "3"	"nein"
[4,] "4"	"ja"



Predictive Maintenance mit MATLAB

- Der Datenanalyse-Schritt des Predictive Maintenance-Prozesses wird mit Hilfe der MATLAB Statistics and Machine Learning Toolbox und des Live Script Editors erstellt.
- Zunächst wird in [Demo 5, Teil 1: MATLAB Live Scripts](#) des elab2go erläutert, wie ein MATLAB Live Script erstellt und verwendet wird.
- Dann wird in [Demo 5, Teil 2: Datenanalyse mit MATLAB](#) die Datenanalyse am Beispiel eines Automobildatensatzes durchgeführt.
- Es wurden dafür zwei Skripte erstellt, ein kommentiertes MATLAB(*.m)-Skript und ein interaktives Live(*.mlx)-Skript, das um visuelle Elementen erweitert wurde.



Datenanalyse mit MATLAB – Die Entwicklungsumgebung

- Die MATLAB Software des US-amerikanischen Unternehmens MathWorks wird in der Industrie und in Forschungseinrichtungen viel für numerische Simulation sowie Datenerfassung, Datenanalyse und -auswertung eingesetzt.
- Für verschiedene Anwendungsgebiete stellt Mathworks Erweiterungspakete bereit, die sogenannten **Toolboxen**.
- Der reine Quellcode, der den Ablauf einer Problemlösung enthält, wird in einer **.m-Datei** gespeichert.



Datenanalyse mit MATLAB – Das Live Script

- In einem Live Script (*.mlx) werden **die Benutzer-Eingaben, der Quellcode selbst, die Ausführung des Quellcodes und die Ausgabe der Ergebnisse** (Grafiken, Werte der Berechnungen) in **einem einzigen interaktiven Fenster** angezeigt.
- Diese Anzeige dient der Übersichtlichkeit und dem Nachvollziehen der Befehle und Funktionen. Welche Auswirkung eine Änderung im Quellcode hat, ist durch die Änderung der ausgegebenen direkt und unmittelbar zu erkennen.
- Durch Ausführung (RUN-Button) oder Änderungen der einzelnen Sektionen ist ein Schritt-für-Schritt Vorgehen bzw. Verständnis möglich.



Datenanalyse mit MATLAB – Der Datenanalyse-Prozess

Die einzelnen Sektionen beinhalten die Schritte:

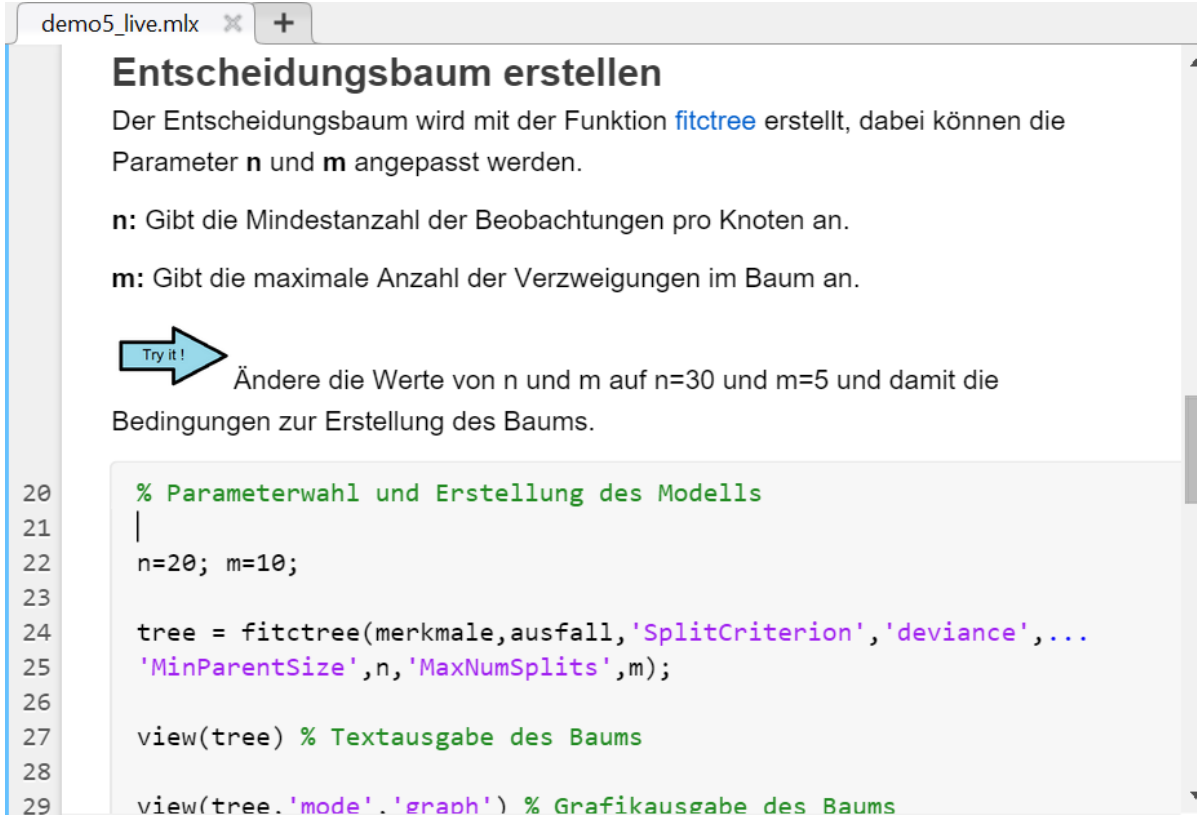
- Einlesen der Trainingsdaten,
- Anzeige verschiedener Merkmale und der Vorhersagevariablen,
- Erstellen des Entscheidungsbaum-Modells anhand von veränderbaren Parametern und
- Verwenden des Modells zur Klassifikation bzw. der Vorhersage eines Ausfalls im Testdatensatz.
- Beurteilung des Modells anhand von Kennzahlen

Die Ausgabe der Sektionen, wenn eine erfolgt, wird in der rechten Spalte neben der Sektion angezeigt. Der Quellcode in den Sektionen ist per Doppelklick auf den Code editierbar.




Datenanalyse mit MATLAB – Der Datenanalyse-Prozess

Die Sektion, in der der Entscheidungsbaum erstellt wird und ein Anpassen der Parameterwerte möglich ist:



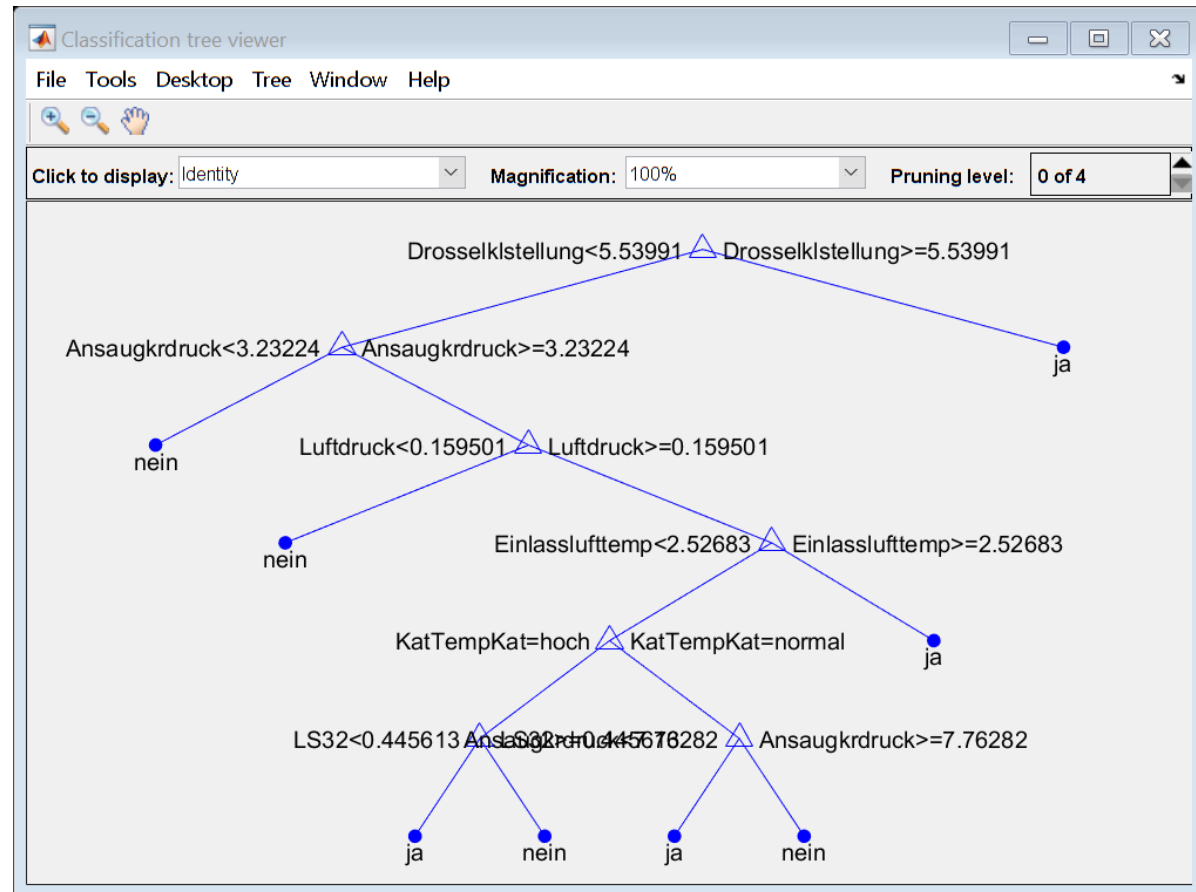
The image shows a MATLAB Live Editor window titled 'demo5_live.mlx'. The main heading is 'Entscheidungsbaum erstellen'. Below it, there is explanatory text about the 'fitctree' function and its parameters 'n' and 'm'. A blue arrow labeled 'Try it!' points to the text 'Ändere die Werte von n und m auf n=30 und m=5 und damit die Bedingungen zur Erstellung des Baums.' Below this, a code block shows the MATLAB code for creating and viewing the decision tree.

```
demo5_live.mlx x +  
Entscheidungsbaum erstellen  
Der Entscheidungsbaum wird mit der Funktion fitctree erstellt, dabei können die  
Parameter n und m angepasst werden.  
  
n: Gibt die Mindestanzahl der Beobachtungen pro Knoten an.  
m: Gibt die maximale Anzahl der Verzweigungen im Baum an.  
  
 Ändere die Werte von n und m auf n=30 und m=5 und damit die  
Bedingungen zur Erstellung des Baums.  
  
20 % Parameterwahl und Erstellung des Modells  
21 |  
22 n=20; m=10;  
23  
24 tree = fitctree(merkmale,ausfall,'SplitCriterion','deviance',...  
25 'MinParentSize',n,'MaxNumSplits',m);  
26  
27 view(tree) % Textausgabe des Baums  
28  
29 view(tree.'mode'. 'graph') % Grafikausgabe des Baums
```



Datenanalyse mit MATLAB – Der Entscheidungsbaum

Der Entscheidungsbaum wählt nur einen Teil der 22 Merkmale zur Vorhersage eines Ausfalls aus.



Datenanalyse mit MATLAB – Der Entscheidungsbaum

- Je höher das Merkmal im Baum steht, d.h. je weiter oben dessen Position ist, desto mehr Einfluss hat es auf die Vorhersage eines Ausfalls.
- In unserem Baum hat die **Drosselklappenstellung den höchsten Einfluss**, da bei einem Messwert größer oder gleich 5.54 direkt ein Ausfall vorhergesagt wird.
- Eine alternative Darstellung des Entscheidungsbaum-Modells ist als Textform gegeben, MATLAB gibt diese Darstellung standardmäßig mit der „view“-Funktion aus.



Datenanalyse mit MATLAB – Der Entscheidungsbaum

Decision tree for classification

```
1  if Drosselklstellung<5.53991 then node 2 elseif Drosselklstellung>=5.53991 then node 3 else ja
2  if Ansaugkrdruck<3.23224 then node 4 elseif Ansaugkrdruck>=3.23224 then node 5 else nein
3  class = ja
4  class = nein
5  if Luftdruck<0.159501 then node 6 elseif Luftdruck>=0.159501 then node 7 else nein
6  class = nein
7  if Einlasslufttemp<2.52683 then node 8 elseif Einlasslufttemp>=2.52683 then node 9 else ja
8  if KatTempKat=hoch then node 10 elseif KatTempKat=normal then node 11 else nein
9  class = ja
10 if LS32<0.445613 then node 12 elseif LS32>=0.445613 then node 13 else nein
11 if Ansaugkrdruck<7.76282 then node 14 elseif Ansaugkrdruck>=7.76282 then node 15 else ja
12 class = ja
13 class = nein
14 class = ja
15 class = nein
```



Datenanalyse mit MATLAB – Die Vorhersage

Die Vorhersage für neue Merkmalskombinationen erfolgt mittels der „predict“-Funktion und wird als ein Vektor ausgegeben:

Vorhersage mittels des Baums

Der erstellte Entscheidungsbaums wird zur Vorhersage eines Ausfall für neue Merkmalskombinationen der Funktion `predict` übergeben.

```
47 disp('Die Vorhersage für die neuen Merkmalskombinationen')
48 disp('bzgl. eines Ausfalls lautet:')
49
50 % Vorhersage mittels Baum
51 vorhersage = predict(tree,merkmale_neu)
```

```
Die Vorhersage für die neuen Merkmalskombinationen
bzgl. eines Ausfalls lautet:
```

```
vorhersage = 4x1 cell array
    {'ja' }
    {'ja' }
    {'nein'}
    {'ja' }
```



Datenanalyse – Zwischenergebnis

- Wir haben wir die Frage nach den Kombinationen von Merkmalen, die zu einem Ausfall eines Motors führen, anhand des Vorhersagemodells des Entscheidungsbaumes beantwortet.
- Dieser Baum wurde danach interpretiert und zur Vorhersage eines Ausfalls bei neu erfassten Daten verwendet.

→ Nun stellt sich aber die Frage wie gut das Vorhersagemodell überhaupt ist.



Datenanalyse im Rahmen der Pred. Maint. – Fortsetzung

- Uns interessiert, wie sicher das Vorhersagemodell den Ausfall eines Motors vorhersagt.
- Anhand der Kennzahlen zum Testfehler lässt sich dies beantworten: Liegen die Kennzahlen in einem vorher festgelegten, akzeptablen Bereich? Wenn ja, dann ist das Modell gut.
- Bei einem guten Modell können wir uns auf die Aussagen und Vorhersagen durch das Modell verlassen.



Datenanalyse – Die Kennzahlen mit MATLAB

```
% Kreuzvalidierung
for i = 1:teilds.NumTestSets
    trIdx = teilds.training(i);
    teIdx = teilds.test(i);
    % Bestimmung des Modells anhand der Trainingsdaten
    treeCV = fitctree(merkmale(trIdx,:),ausfall(trIdx,:), 'SplitCriterion','deviance',...
        'MinParentSize',n, 'MaxNumSplits',m);

    vorhersageCV = predict(treeCV,merkmale(teIdx,:));
```

- Der Kennzahl-Sektion besteht aus einer for-Schleife, die bei jedem Schleifen-Durchlauf einen Teildatensatz auswählt, mit den übrigen Teildatensätzen das Vorhersagemodell/Baum erstellt und dann für ausgewählten Teildatensatz eine Vorhersage durchführt. Die 10 Teildatensätze werden mit der Funktion `cvpartition` erstellt.

```
anzTeil = 10;
teilds = cvpartition(x, 'k',anzTeil);
```



Datenanalyse – Die Kennzahlen mit MATLAB

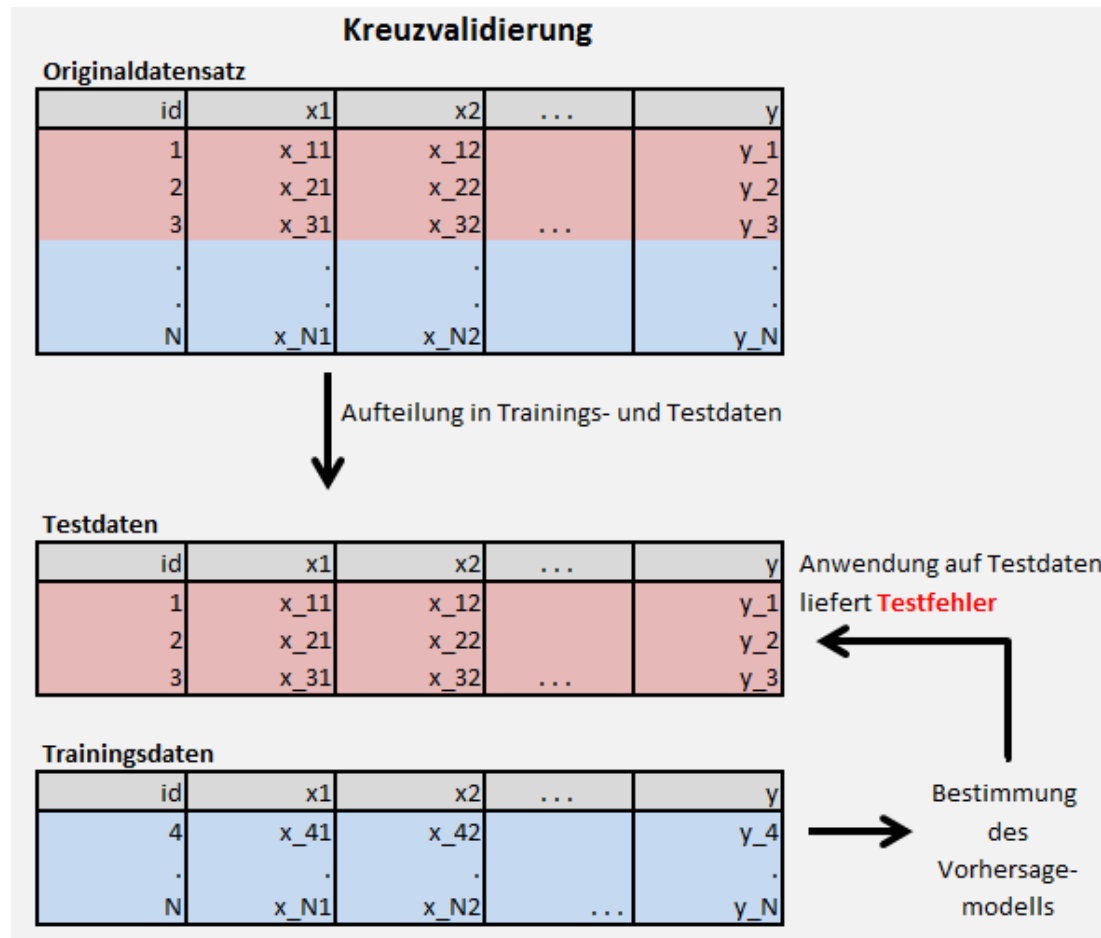
- Die vorhergesagten Werte werden dann mit der Funktion `cfmatrix2` mit den tatsächlichen Werten (actual) für die Variable Ausfall verglichen und somit die Kennzahlen bestimmt.

```
cf = cfmatrix2(actual, vorhersageCV, [0,1], 0, 0);  
|
```

- Dies erfolgt K mal und nachdem die Schleife durchlaufen wurde, werden alle K Werte der Kennzahlen gemittelt.
→ dies entspricht der Idee der K-fachen Kreuzvalidierung/CrossValidation.



Datenanalyse – Die Kennzahlen mit MATLAB



Datenanalyse – Die Beurteilung des Modells

Wie gut der Entscheidungsbaum zur Vorhersage eines Ausfalls geeignet ist, können wir nun anhand der Kennzahlen zum Testfehler beurteilen:

<u>Precision</u>	<u>Recall</u>	<u>Accuracy</u>
0.66063	0.81925	0.66209

Bei einem „guten“ Modell liegen die Kennzahlen „nahe“ an 100%.

-----Beurteilung des Modells-----

Das Modell, bewertet mit Precision, ist moderat.
 Das Modell, bewertet mit Recall, ist gut.
 Das Modell, bewertet mit Accuracy, ist moderat.



Zusammenfassung

- Die Vorausschauende Wartung ist ein Prozess aus 5 Schritten, einer davon ist die **Datenanalyse**.
- In der Datenanalyse werden durch Anwendung von Algorithmen und Verfahren des maschinellen Lernens Modelle zur **Vorhersage eines Ausfalls** einer Maschine oder eines Bauteils erstellt.
- Die Datenanalyse beinhaltet auch die **Validierung dieser Modelle** um die Sicherheit eines Ausfalls zu bestimmen.
- Durch eine langfristige Datenerhebung können die Vorhersagemodelle verbessert und gezielt auf ein Bauteil ausgerichtet werden. (→ **Updates**)



Zusammenfassung

- Beispielhaft wird im elab2go die Datenanalyse anhand eines Automobildatensatzes durchgeführt.
- **Verschiedene Tools (RapidMiner, RStudio, MATLAB)** stellen die zur Datenanalyse benötigten Algorithmen und Verfahren zur Verfügung und je nach Kontext kann eine Auswahl eines Tools erfolgen.

→ Durch die Predictive Maintenance wird **frühzeitig** vor dem Ausfall einer Maschine/Bauteils gewarnt und somit **Ausfallzeiten und Kosten niedrig** gehalten.

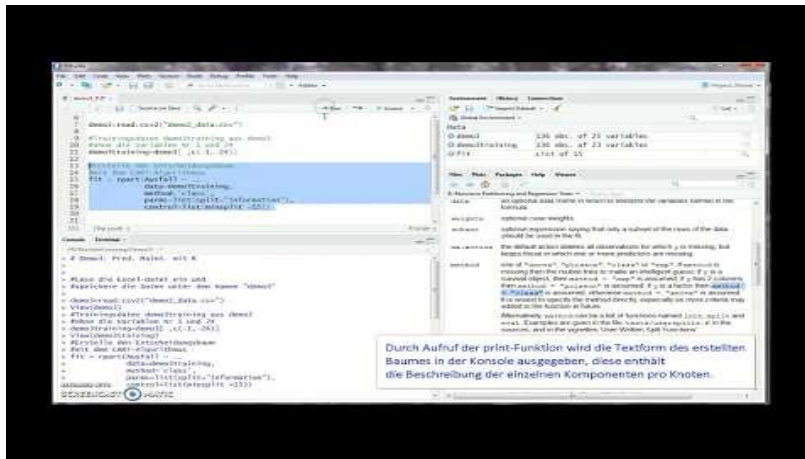


Weitere Informationen

- Wir stellen regelmäßig Lernvideos auf unserem [YouTube-Kanal](#) ein.



elab2go
Mobile
Engineering
Lab



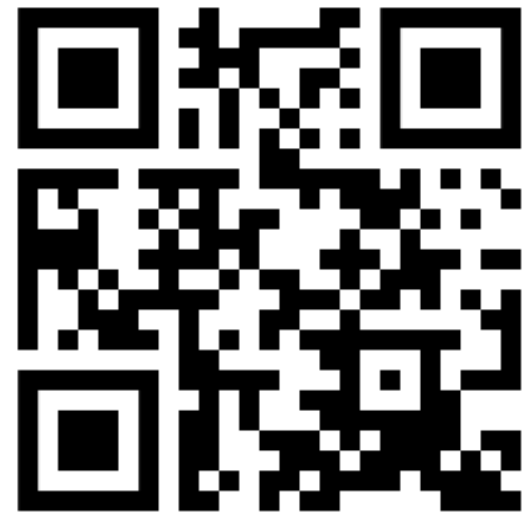
Weitere Informationen

Kontakt elab2go Mobile Engineering Lab

Prof. Dr. Eva Maria Kiss, Studiengangsleiterin Elektrotechnik
M. Sc. Anke Welz, Offene Digitalisierungsallianz Pfalz

Hochschule Kaiserslautern
Fachbereich Angewandte Ingenieurwissenschaften
Schoenstrasse 11, 67659 Kaiserslautern

E-Mail:
evamaria.kiss(at)hs-kl.de
anke.welz(at)hs-kl.de



www.elab2go.de



**VIELEN DANK FÜR IHRE
AUFMERKSAMKEIT!**

