



SymReg

JOSEF RESSEL CENTER FOR
SYMBOLIC REGRESSION



Welchen Beitrag kann symbolische Regression zu Explainable AI leisten? – Praxis-Beispiele aus dem Josef Ressel Zentrum

Softwarepark Hagenberg IT-Expert*innenreihe „Thinking AI“

FH-Prof. DI Dr. Gabriel Kronberger

17. Februar 2022

<https://symreg.at>



Wozu „erklärbare KI“?

- Zur Sicherstellung, dass sich die KI richtig verhält.
 - Wenn wir von der KI etwas lernen wollen.
 - Zur Erkennung von Fehlern im Modell oder -Algorithmus.
- Erklärbarkeit ist „nice-to-have“ aber nicht immer zwingend notwendig.

Github CoPilot

<https://copilot.github.com/>

```
1 package main
2
3 type Run struct {
4     Time int // in milliseconds
5     Results string
6     Failed bool
7 }
8
9 // Get average runtime of successful runs in seconds
10 func averageRuntimeInSeconds|
11
12
13
14
15
16
17
18
19
20
21
22
23
```

Github CoPilot

<https://copilot.github.com/>

```
1 package main
2
3 type Run struct {
4     Time int // in milliseconds
5     Results string
6     Failed bool
7 }
8
9 // Get average runtime of successful runs in seconds
10 func averageRuntimeInSeconds(runs []Run) float64 {
11     var totalTime int
12     var failedRuns int
13     for _, run := range runs {
14         if run.Failed {
15             failedRuns++
16         } else {
17             totalTime += run.Time
18         }
19     }
20
21     averageRuntime := float64(totalTime) / float64(len(runs) - failedRuns) / 1000
22     return averageRuntime
23 }
```

AlphaCode

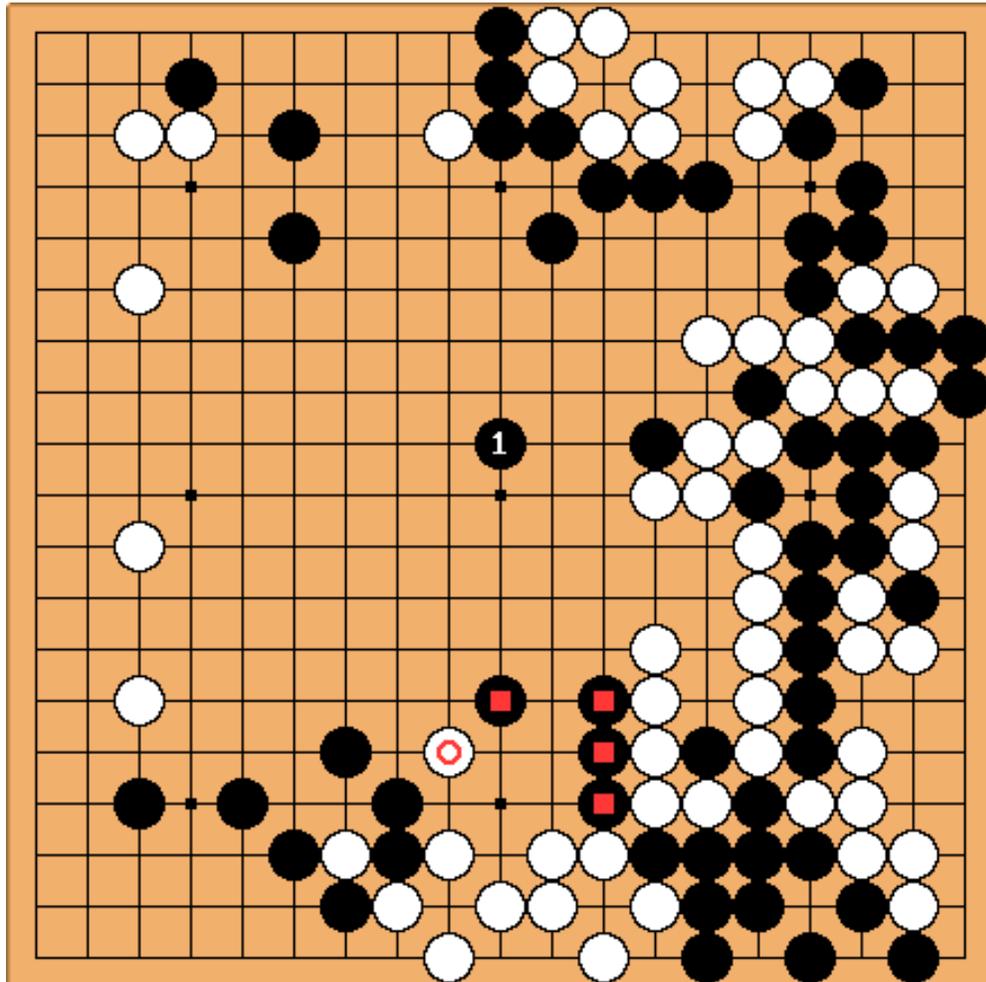
Problem Description

```
# RATING: 1200↵
# TAGS: *special↵
# LANGUAGE IS python3↵
# CORRECT SOLUTION↵
# Let's define  $S(x)$  to be the sum of digits of number  $x$  written in decimal↵
# system. For example,  $S(5) = 5$ ,  $S(10) = 1$ ,  $S(322) = 7$ .↵
# ↵
# We will call an integer  $x$  interesting if  $S(x + 1) < S(x)$ . In each test you↵
# will be given one integer  $n$ . Your task is to calculate the number of integers↵
#  $x$  such that  $1 \leq x \leq n$  and  $x$  is interesting.↵
# ↵
# Input↵
# ↵
# The first line contains one integer  $t$  ( $1 \leq t \leq 1000$ ) – number of test cases.↵
# ↵
# Then  $t$  lines follow, the  $i$ -th line contains one integer  $n$  ( $1 \leq n \leq 10^9$ ) for↵
# the  $i$ -th test case.↵
# ↵
# Output↵
# ↵
# Print  $t$  integers, the  $i$ -th should be the answer for the  $i$ -th test case.↵
# ↵
```

Solution

```
t = int(input())↵
↵
while t > 0:↵
    n = int(input())↵
    s = 0↵
    ↵
    for i in range(n+1):↵
        s += sum(map(int, str(i)))↵
        if s > 9:↵
            break↵
    ↵
    print(n // 10 + (n % 10 == 9))↵
    t -= 1↵
```

<https://alphacode.deepmind.com/>



Menschliche Erklärung:

- **1** erweitert das schwarze Moyo oben,
- hilft den vier markierten schwarzen Steinen,
- reduziert den Einfluss der stabilen weißen Position rechts
- und bereitet eine Reduktion des weißen Moyos auf der linken Seite vor.

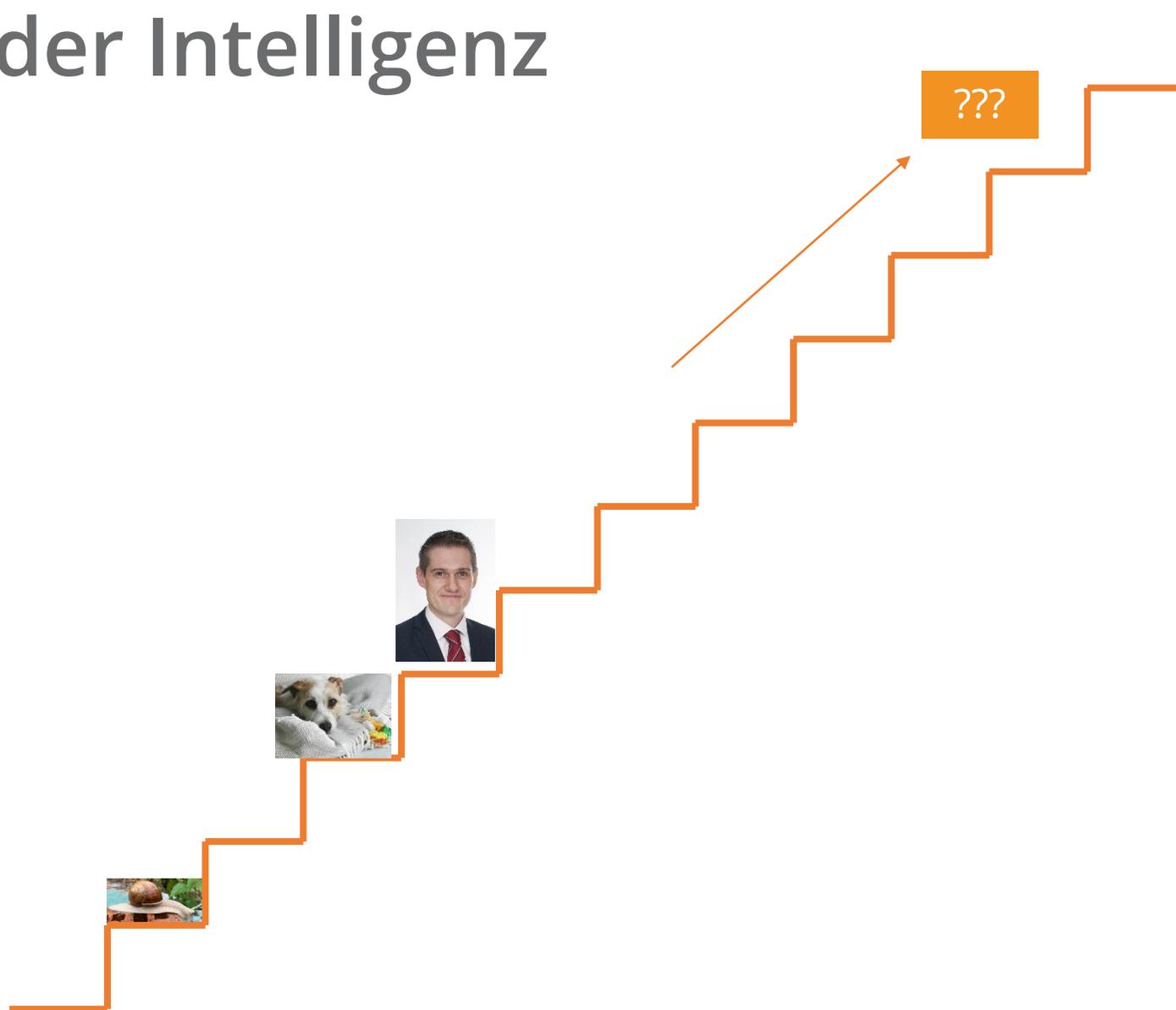
<https://senseis.xmp.net/?EarReddeningMove>

Stufen der Intelligenz



nach Tim Urban,
waitbutwhy.com

Stufen der Intelligenz



nach Tim Urban,
waitbutwhy.com

Was ist symbolische Regression?

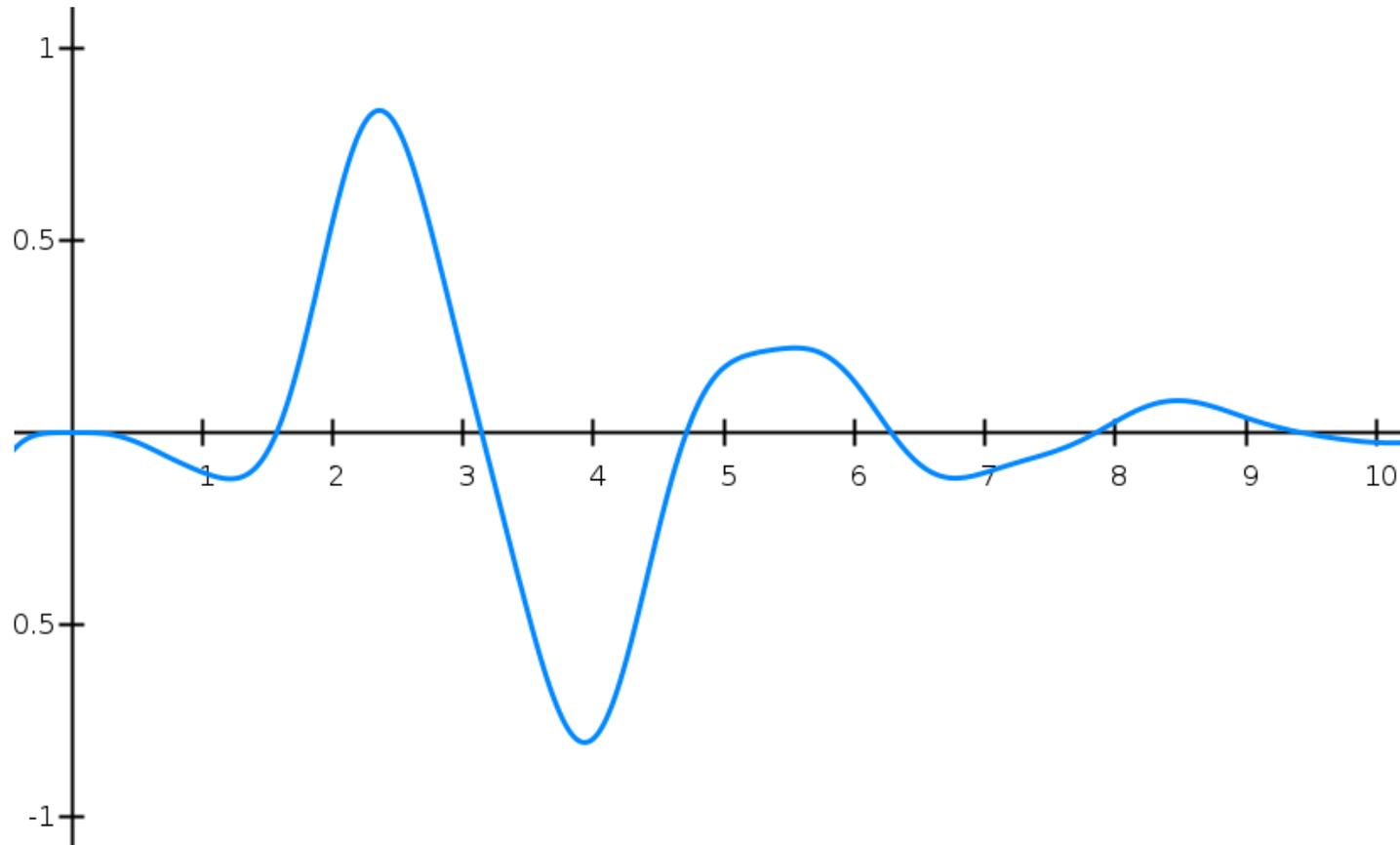
Vereinfachte Beschreibung:

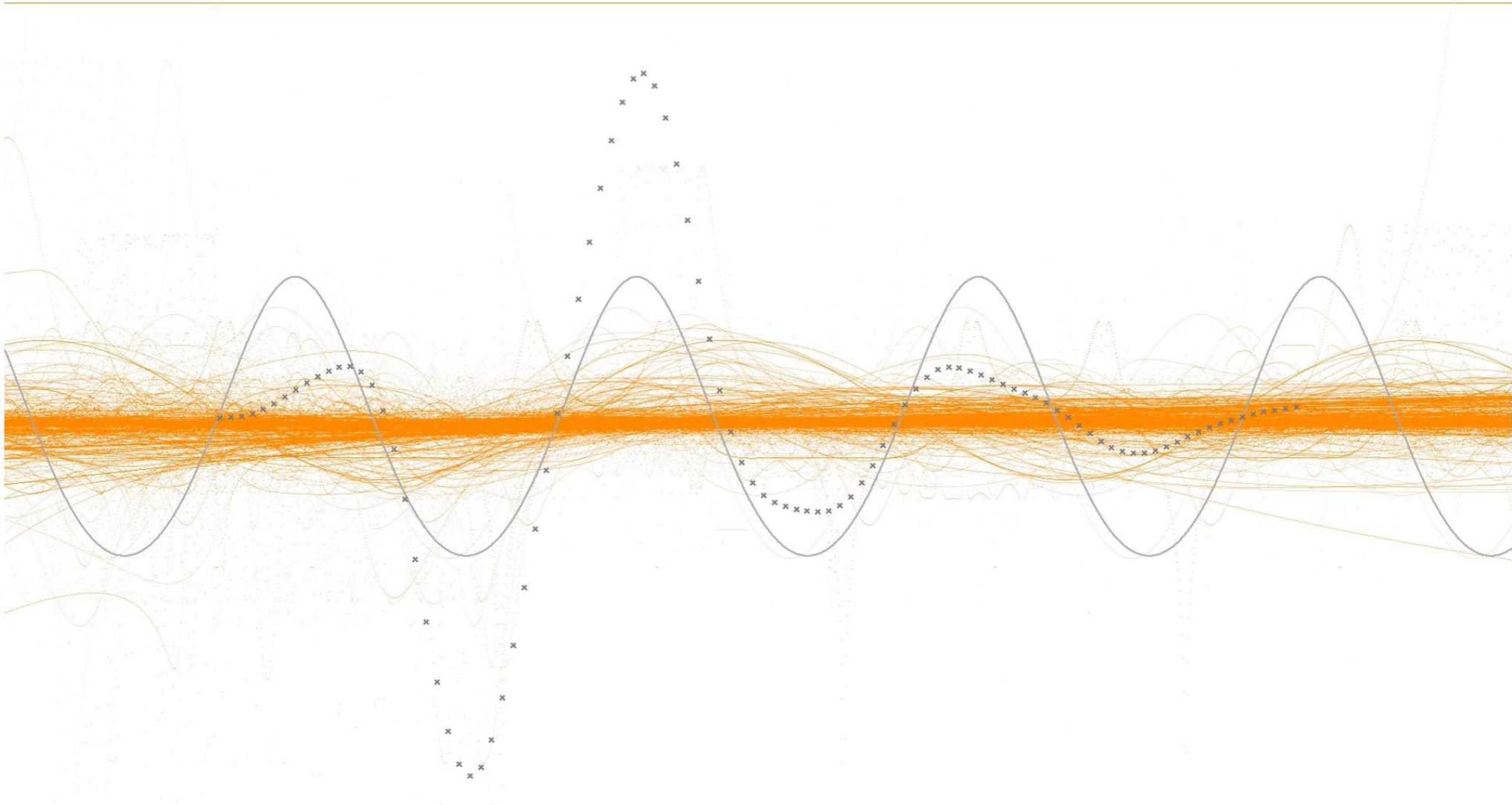
Suche nach einer Formel, die zu Daten passt.

Eine fundamentale Problemstellung in vielen wissenschaftlichen und technischen Disziplinen.

Was ist symbolische Regression?

Bsp.:





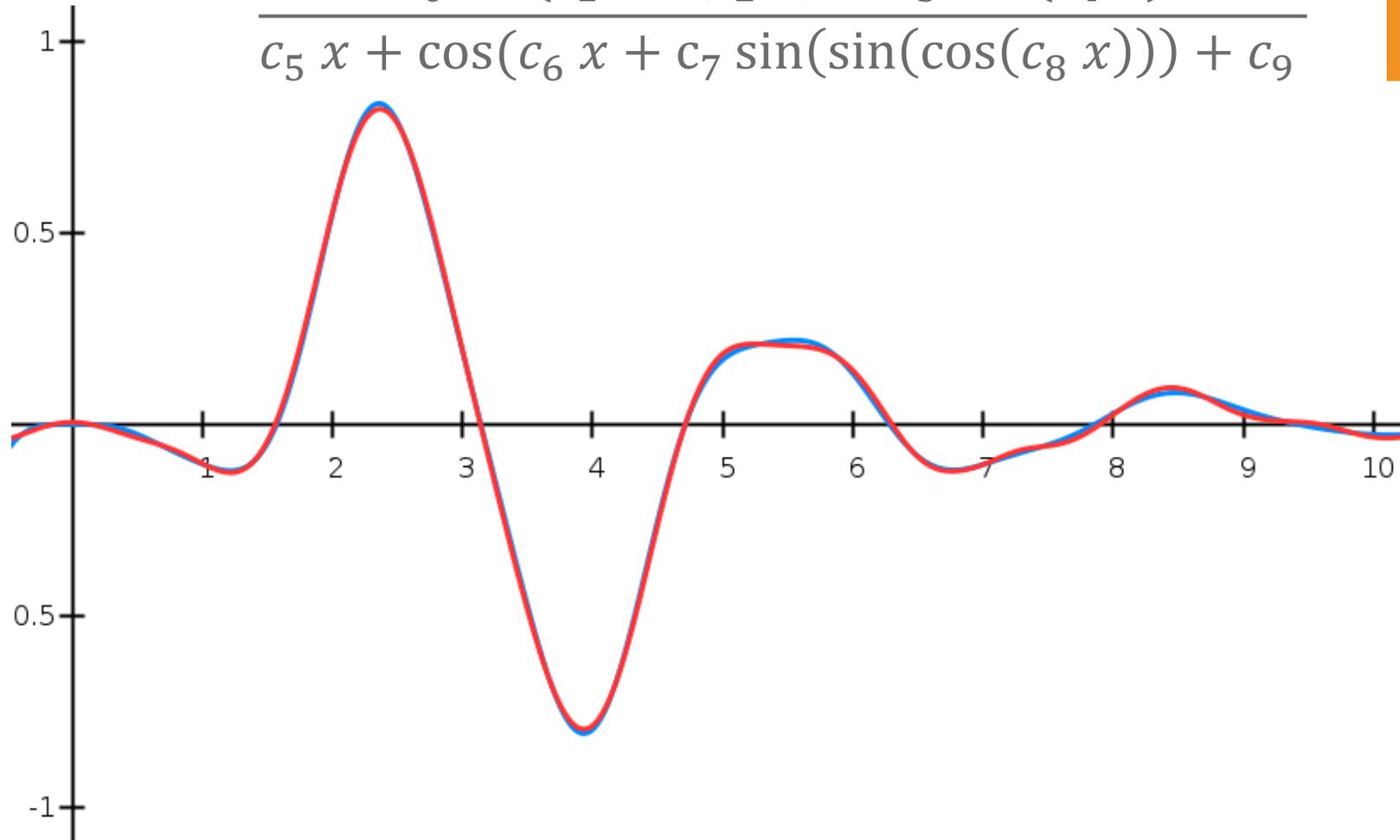
$\text{EXP}(\text{COS}(\text{SIN}(\text{((-1}^*X) + \text{COS}(\text{COS}(\text{SIN}(\text{SIN}(\text{COS}(\text{LOG}(\text{((NaN}^*X) + (\text{NaN})) / ((-1}^*X) + 6.3))))))))))$

<https://symreg.at>

Was ist symbolische Regression?

$$\frac{c_0 \sin(c_1 \cos(c_2 x)) + c_3 \cos(c_4 x)}{c_5 x + \cos(c_6 x + c_7 \sin(\sin(\cos(c_8 x))))} + c_9$$

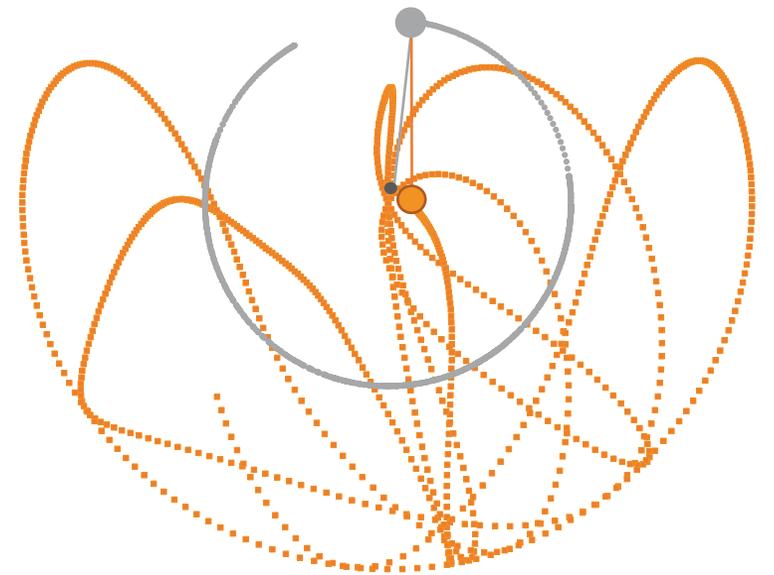
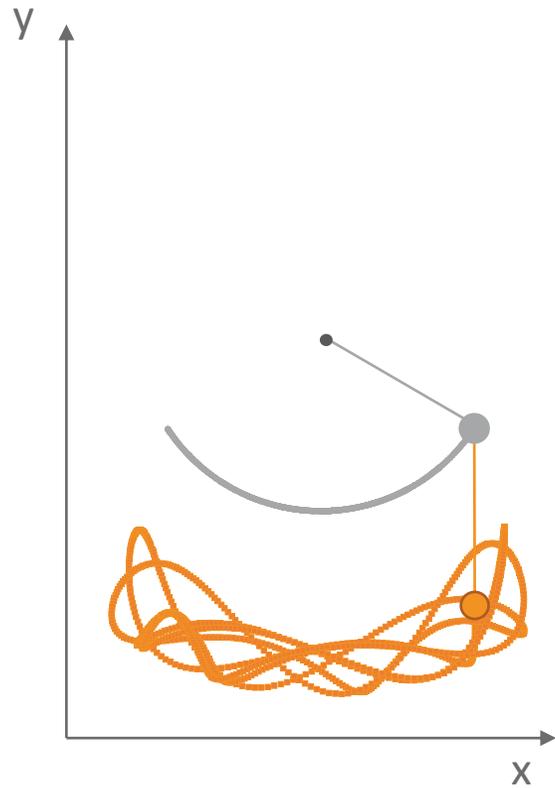
Direkt
erklärbar!



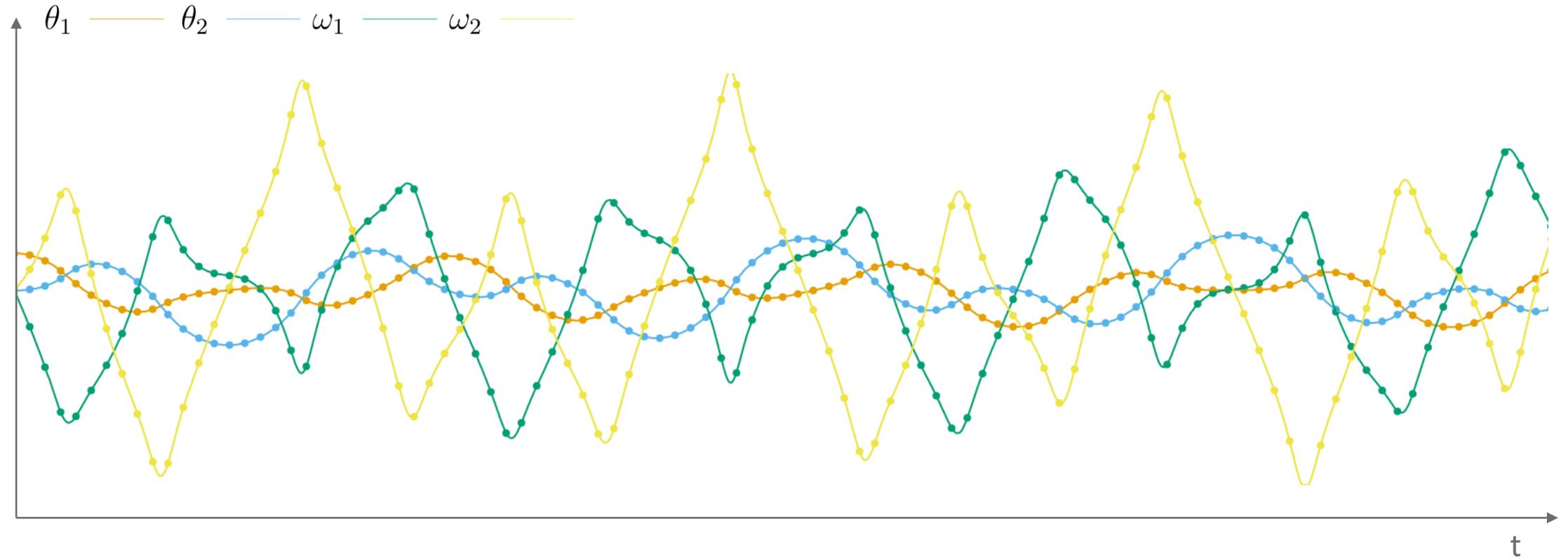
Symbolische Regression und Physik

- Mathematik ist die Sprache der Physik.
- Physik ist die Lehre der Gesetze des Universums.
- Viele technologische Fortschritte werden durch besseres Verständnis der Physik ermöglicht.
- Fundament: **Formeln und Simulation**

Mechanik: Doppelpendel



Mechanik: Doppelpendel



Idealisiertes physikalisches Modell

nach **Newton**: $F = m a$.

Daraus lässt sich das gewöhnliche Differenzialgleichungssystem für das Doppelpendel herleiten:

$$\ddot{\theta}_1 = (l_1 s)^{-1} (m_2 g \sin \theta_2 \cos(\theta_1 - \theta_2) - m_2 \sin(\theta_1 - \theta_2) (l_1 \dot{\theta}_1^2 \cos(\theta_1 - \theta_2) + l_2 \dot{\theta}_2^2) - (m_1 + m_2) g \sin \theta_1)$$

$$\ddot{\theta}_2 = (l_2 s)^{-1} ((m_1 + m_2) (l_1 \dot{\theta}_1^2 \sin(\theta_1 - \theta_2) - g \sin \theta_1 \cos(\theta_1 - \theta_2)) + m_2 l_2 \dot{\theta}_2^2 \sin(\theta_1 - \theta_2) \cos(\theta_1 - \theta_2))$$

$$s = m_1 + m_2 \sin(\theta_1 - \theta_2)^2$$

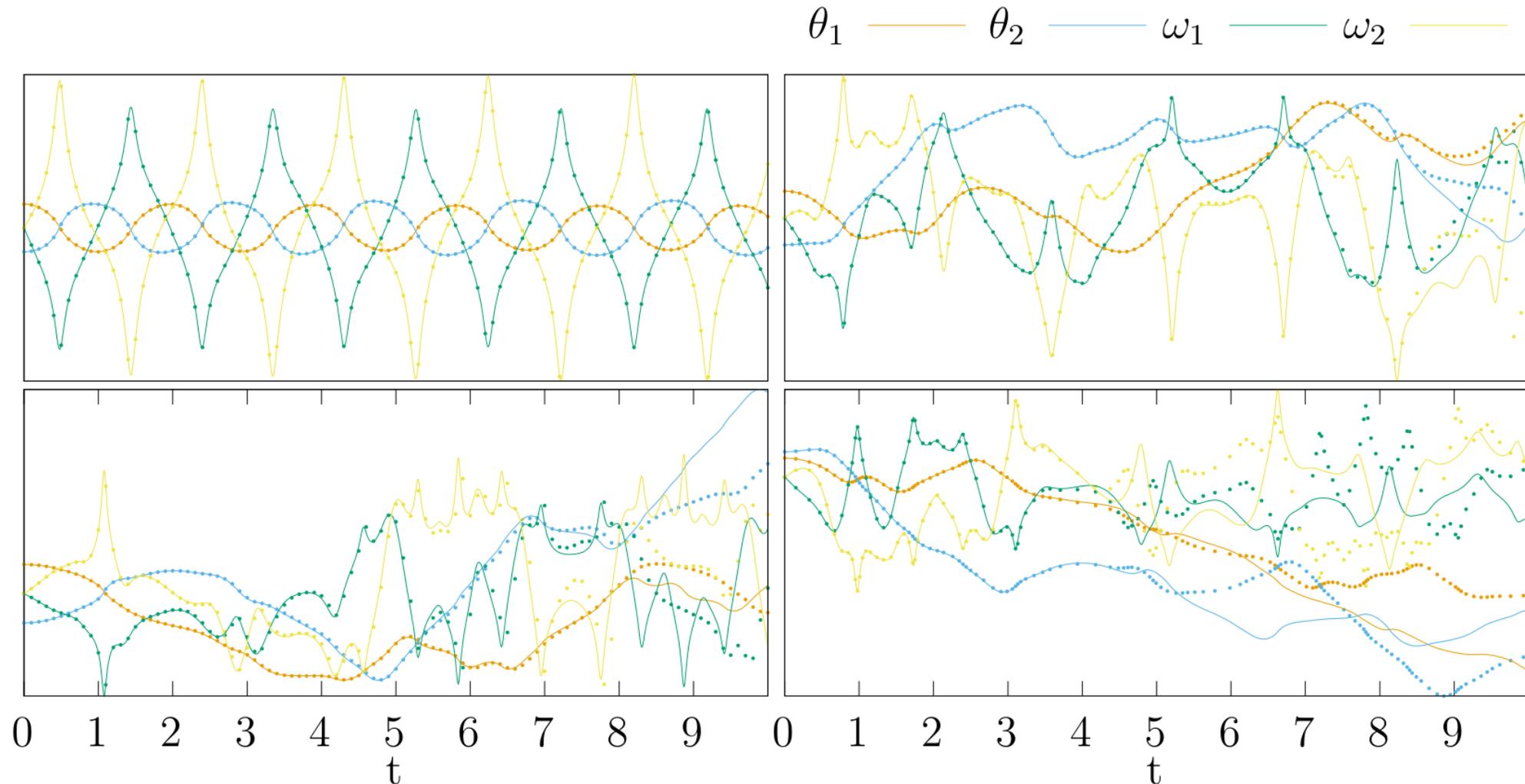
SymReg Modell

Von der KI ausschließlich aus Beobachtung gefundenes Gleichungssystem!

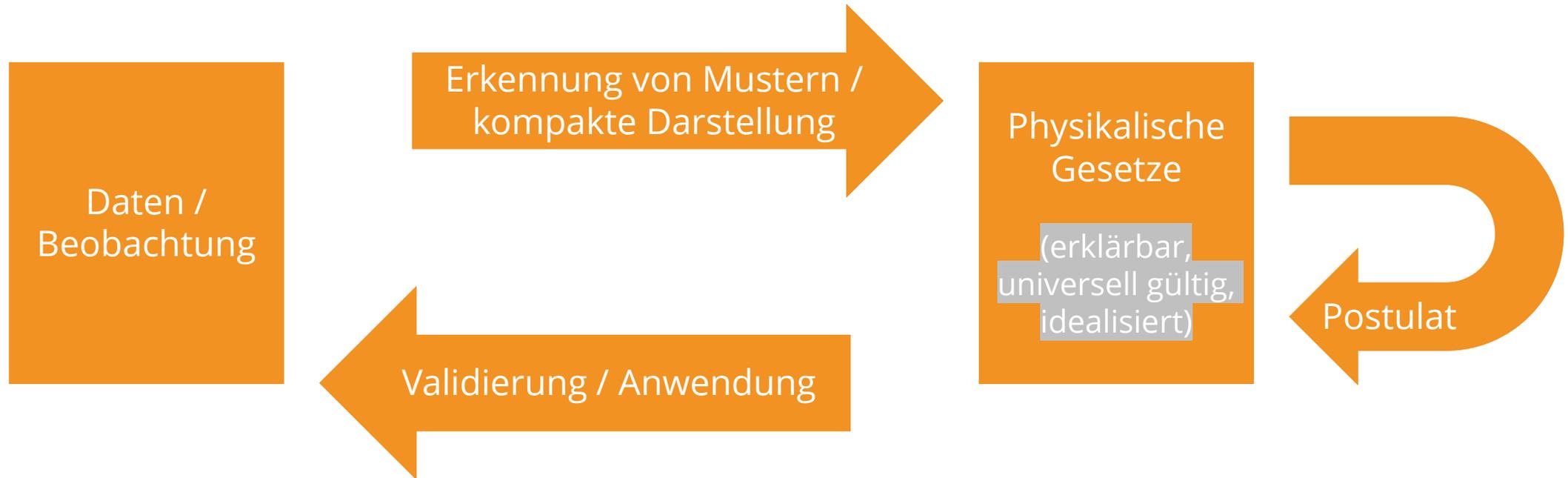
$$\ddot{\theta}_1 = -\sin(0.6719 \sin(1.851 \sin(\sin(\theta_2 - \theta_1)))) \\ \left(-0.8034 \dot{\theta}_2^2 + \cos(\theta_2 - \theta_1) \left(-0.8034 \dot{\theta}_1^2 - 7.882 \sin(\theta_1 + 14.13)\right) + 1.757 \times 10^{-7}\right) \\ + 9.81 \sin(\theta_1 + 15.7) - 2.322 \times 10^{-5}$$

$$\ddot{\theta}_2 = \sin(\sin \sin \sin \sin(\theta_2 - \theta_1) + \sin \sin(\theta_2 - \theta_1)) \\ \left(-0.5013 \cos(\theta_2 - \theta_1) \dot{\theta}_2^2 - 1.002 \dot{\theta}_1^2 - 9.842 \cos \theta_1\right) + 5.066 \times 10^{-4}$$

Simulation des SymReg Modells für unbekannte Anfangszustände

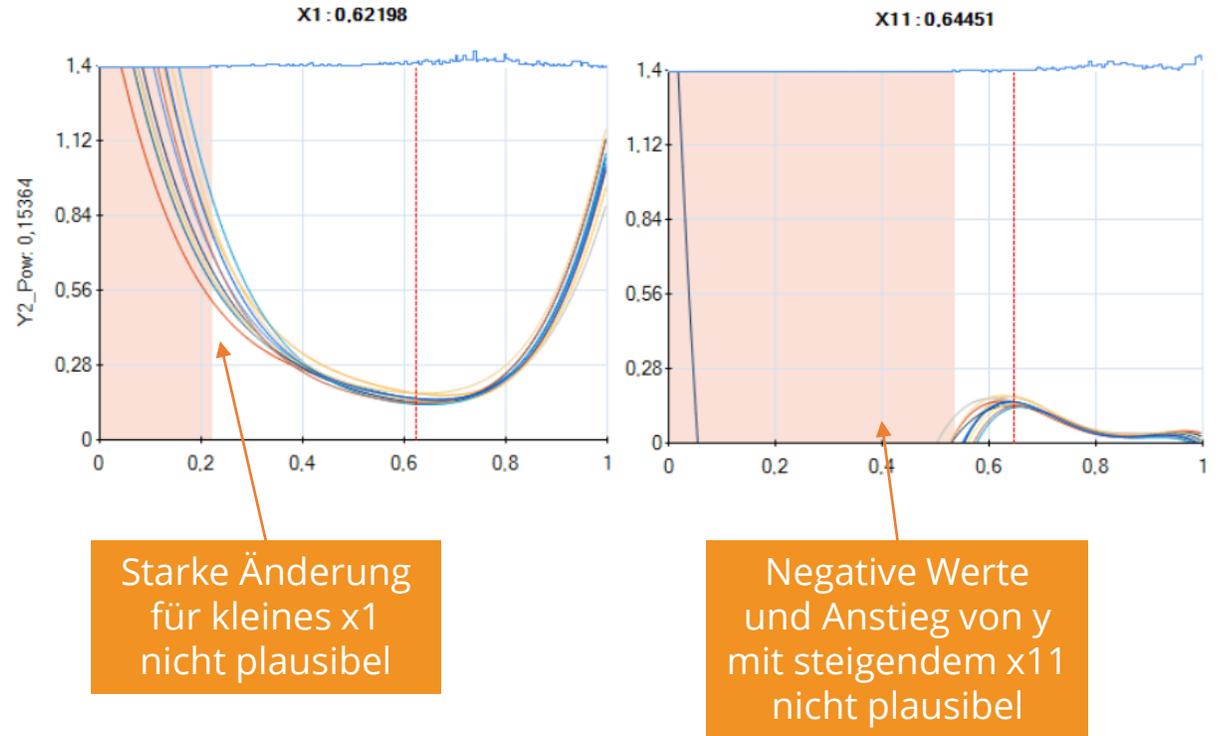
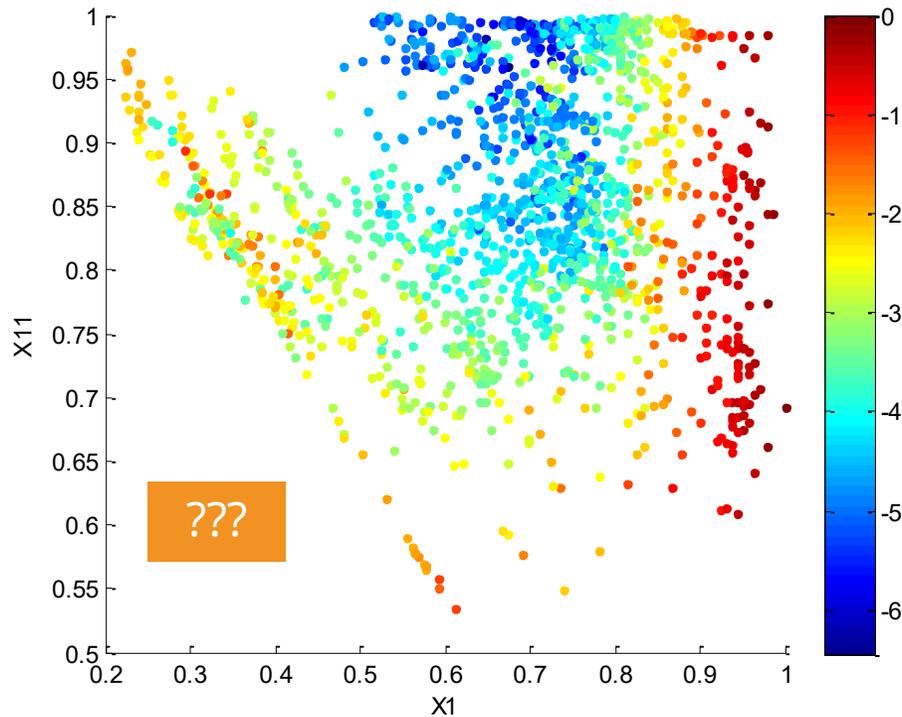


Der ewige Zyklus der Erkenntnis

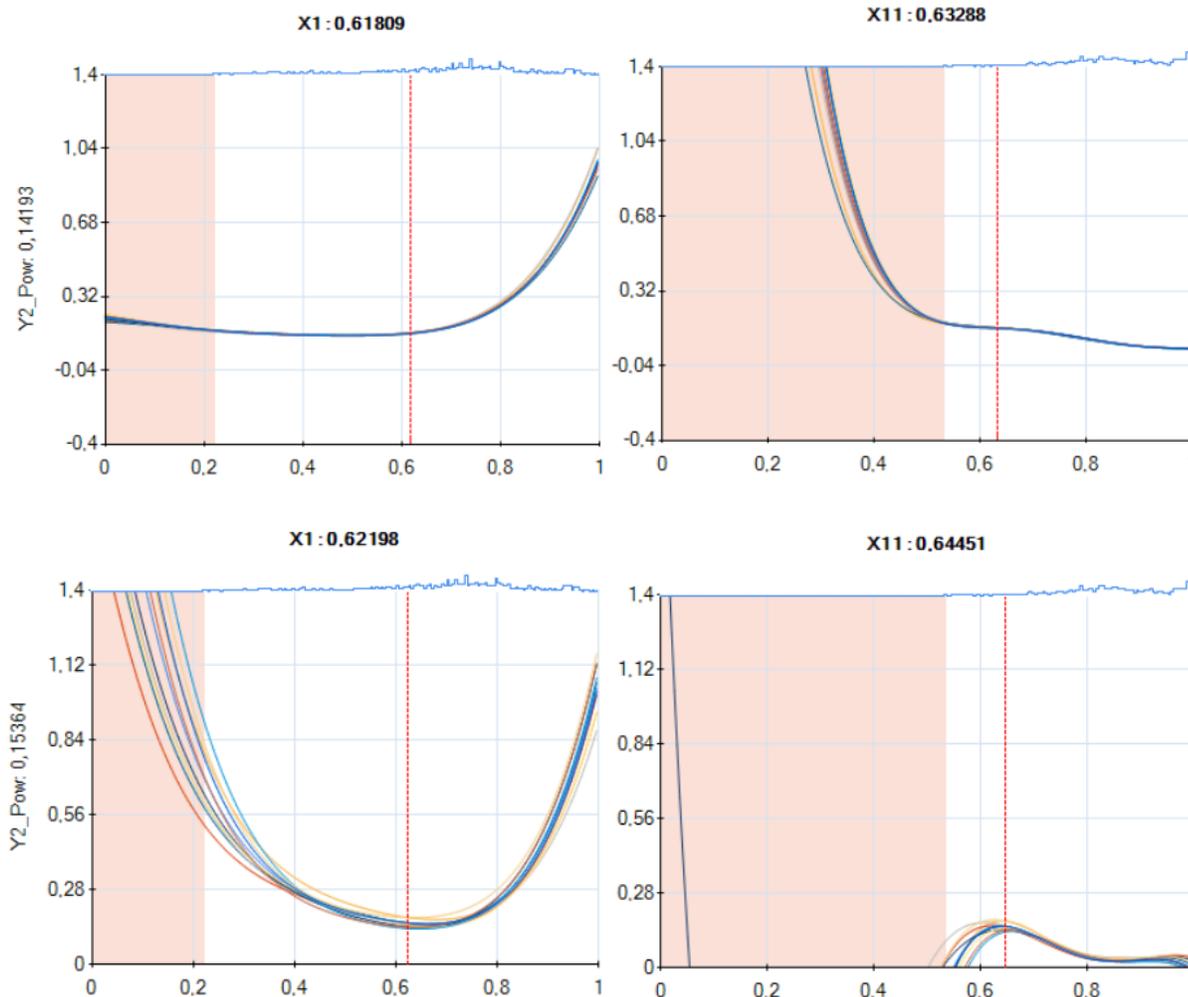


Prinzipiell vollständig durch KI automatisierbar.
Es wäre dumm sich ausschließlich auf Daten zu verlassen.

Extrapolation – Ein Problem für rein Daten-basierte KI-Ansätze



Verbesserung der Extrapolation durch Einschränkungen



Anpassung auf Daten & physikalisch motivierte Einschränkungen

$$y = P(x_1, x_2, x_3, \dots, x_{11}) + \epsilon$$

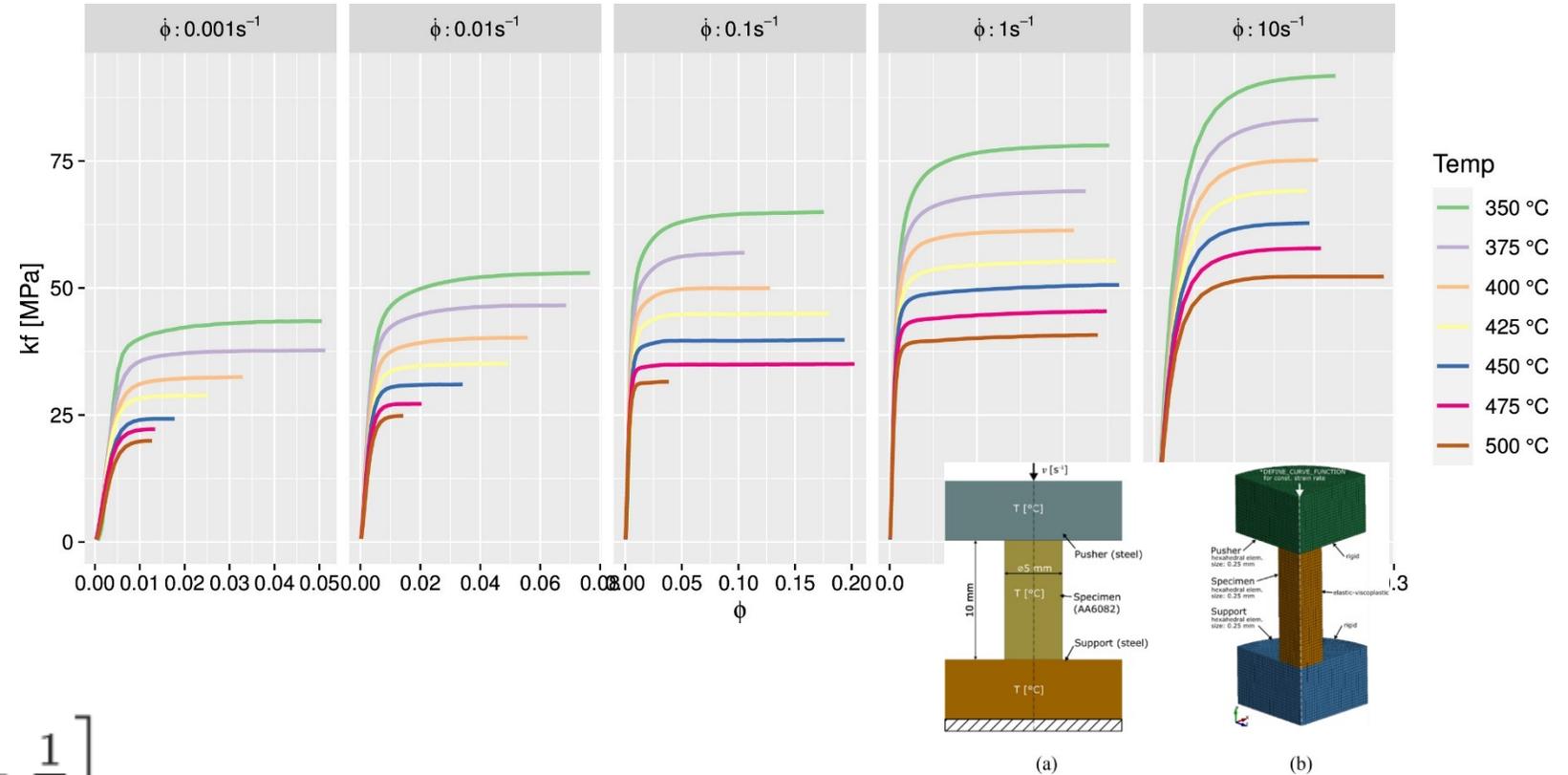
$$P \geq 0, \quad \frac{\partial P}{\partial x_1} \geq -0.2, \quad \frac{\partial P}{\partial x_{11}} \leq 0,$$

$$x_1, x_{11} \in [0..1]$$

Zum Vergleich ohne Einschränkung

Erweiterung von physikalischen Modellen

Spannungs-Dehnungs-Kurven



Physikalisches
Versetzungsdichte-Modell:

$$\sigma = \sigma_y + MGb \left[\frac{\sqrt{\rho}}{2} + \frac{1}{\delta} \right]$$

$$d\rho = \frac{M\sqrt{\rho}}{bA} \dot{\varphi} dt - 2BM \frac{d_{ann}}{b} \rho M \dot{\varphi} dt - 2CD \frac{Gb^3}{k_B T} [\rho^2 - \rho_{eq}^2] dt$$

[Kablman et al., 2022]

Erweiterung von physikalischen Modellen

[Kronberger et al., 2022]

$$\sigma = \sigma_y + M G b \left[\frac{\sqrt{\rho}}{2} + \frac{1}{\delta} \right]$$

Physik

$$d\rho = \frac{M\sqrt{\rho}}{bA} \dot{\varphi} dt - 2BM \frac{d_{ann}}{b} \rho M \dot{\varphi} dt - 2CD \frac{Gb^3}{k_B T} [\rho^2 - \rho_{eq}^2] dt$$

$$u(T, \dot{\varphi}) = -\log(0.11 T + 3.734 \dot{\varphi} - 17.34) - 0.069 \log(\dot{\varphi})$$

$$v(T) = -4.35 \log(T) + 4.938$$

$$w(T, \dot{\varphi}) = -13.708 \log(\dot{\varphi}) T^{-1} - 10205 \log(\dot{\varphi}) T^{-2} + 13675 T^{-2} \\ + 0.777 \log(\dot{\varphi}) - 0.8657$$

Symbolische Regression

Was kann symbolische Regression beitragen?

- Kann erklärbare Formeln liefern (im Idealfall sogar mit physikalisch interpretierbaren Parametern).
- Ist speziell geeignet für die Modellierung technischer / physikalische Systeme.
- Bietet interessante Möglichkeiten speziell in Kombination mit physikalischen Modellen.

Literaturverweise

[Kronberger et al., 2021] Kronberger, G., de França, F. O., Burlacu, B., Haider, C., & Kommenda, M. (2021). Shape-constrained symbolic regression–improving extrapolation with prior knowledge. *Evolutionary computation*, 1-24.

[Kabliman et al., 2021] Kabliman, E., Kolody, A. H., Kronsteiner, J., Kommenda, M., & Kronberger, G. (2021). Application of symbolic regression for constitutive modeling of plastic deformation. *Applications in Engineering Science*, 6, 100052.

[Kronberger et al., 2022] Kronberger, G., Kabliman, E., Kronsteiner, J., & Kommenda, M. (2022). Extending a physics-based constitutive model using genetic programming. *Applications in Engineering Science*, 9, 100080.



SymReg

JOSEF RESSEL CENTER FOR
SYMBOLIC REGRESSION

<https://symreg.at>



Welchen Beitrag kann symbolische Regression zu Explainable AI leisten? – Praxis-Beispiele aus dem Josef Ressel Zentrum

Softwarepark Hagenberg IT-Expert*innenreihe „Thinking AI“

FH-Prof. DI Dr. Gabriel Kronberger

17. Februar 2022