



Algorithmus Schmiede

Data Science | Numerik | Physik

Data Science und Machine Learning für Unternehmen

Dr. Markus Dutschke

Frühjahrskonferenz der Wirtschaftsunioren Deutschland
Wetzlar, April 2023



AI Tools für Unternehmer: ChatGPT



Schreibe mir eine SEO optimierten Websitext zum Keyword "Data Science für Unternehmen".



Gerne! Hier ist ein SEO-optimierter Websitext zum Keyword "Data Science für Unternehmen":

Data Science für Unternehmen - wie Sie durch Datenanalyse Ihre Geschäftsprozesse optimieren können

In der heutigen Geschäftswelt sind Daten der Schlüssel zum Erfolg. Doch wie können Unternehmen die riesigen Mengen an Daten, die sie täglich generieren, effektiv nutzen? Die Antwort lautet: Data Science.

Data Science ist ein interdisziplinärer Ansatz, der Statistik, Programmierung und maschinelles Lernen kombiniert, um Erkenntnisse aus Daten zu gewinnen. Unternehmen, die Data Science in ihre Geschäftsprozesse integrieren, können ihre Entscheidungsfindung verbessern, Prozesse automatisieren, Kosten senken und ihre Wettbewerbsfähigkeit stärken.

Unser Unternehmen bietet Ihnen maßgeschneiderte Lösungen für Ihre Data-Science-Bedürfnisse. Wir arbeiten eng mit Ihnen zusammen, um Ihre Datenquellen zu identifizieren, Datenmodelle zu entwickeln und Erkenntnisse zu gewinnen, die für Ihre Geschäftsziele relevant sind.

Unsere Data-Science-Lösungen umfassen:

- Datenanalyse und -visualisierung
- Vorhersagemodellierung
- Anomalieerkennung
- Textanalyse
- und vieles mehr

Unsere Experten haben jahrelange Erfahrung in der Anwendung von Data Science in verschiedenen Branchen, darunter Finanzen, Gesundheitswesen, Marketing und E-Commerce. Wir sind stolz darauf, unseren Kunden maßgeschneiderte Lösungen zu bieten, die auf ihre spezifischen Bedürfnisse zugeschnitten sind.

Kontaktieren Sie uns noch heute, um mehr darüber zu erfahren, wie Data Science Ihr Unternehmen transformieren kann.

AI Tools für Unternehmer: Midjourney, YOLO

Midjourney
midjourney.com

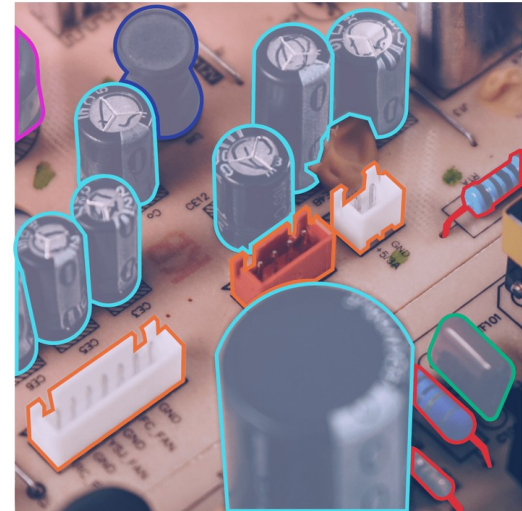
"group of business people walking through giant amazon warehouse, photorealistic"



Lizenz: <https://creativecommons.org/licenses/by-nc/4.0/>

YOLO v8
ultralytics.com/yolov8

Segmentation

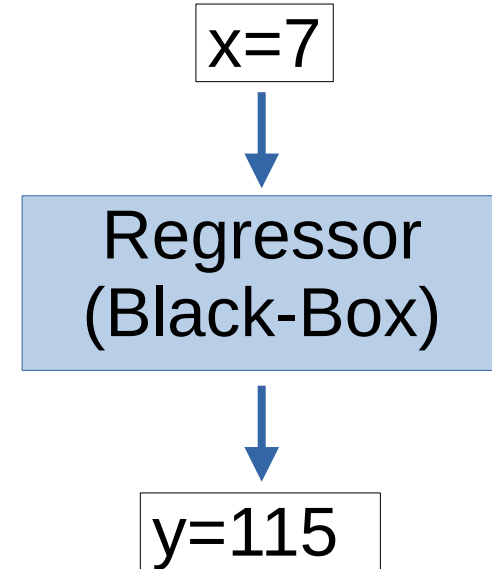


Capacitor, Resistor, Transformer,
Connector, Inductor, Polyester Capacitor

zu den Basics: Was ist ein Regressor?

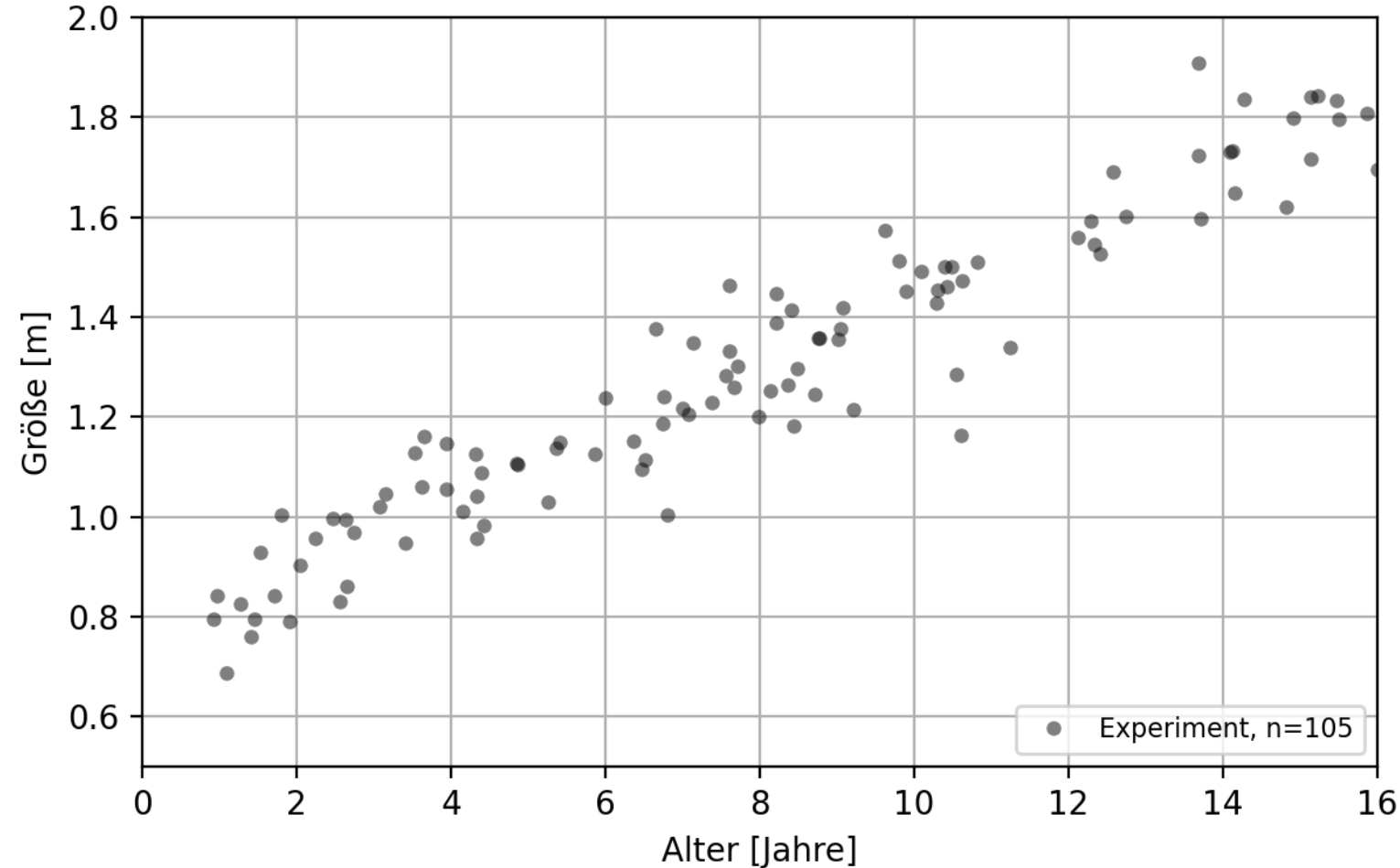
Einsatz:

- x-Wert hineinwerfen
y-Wert herausbekommen
- Training mit x-y Paaren
statt Implementierung von Formel



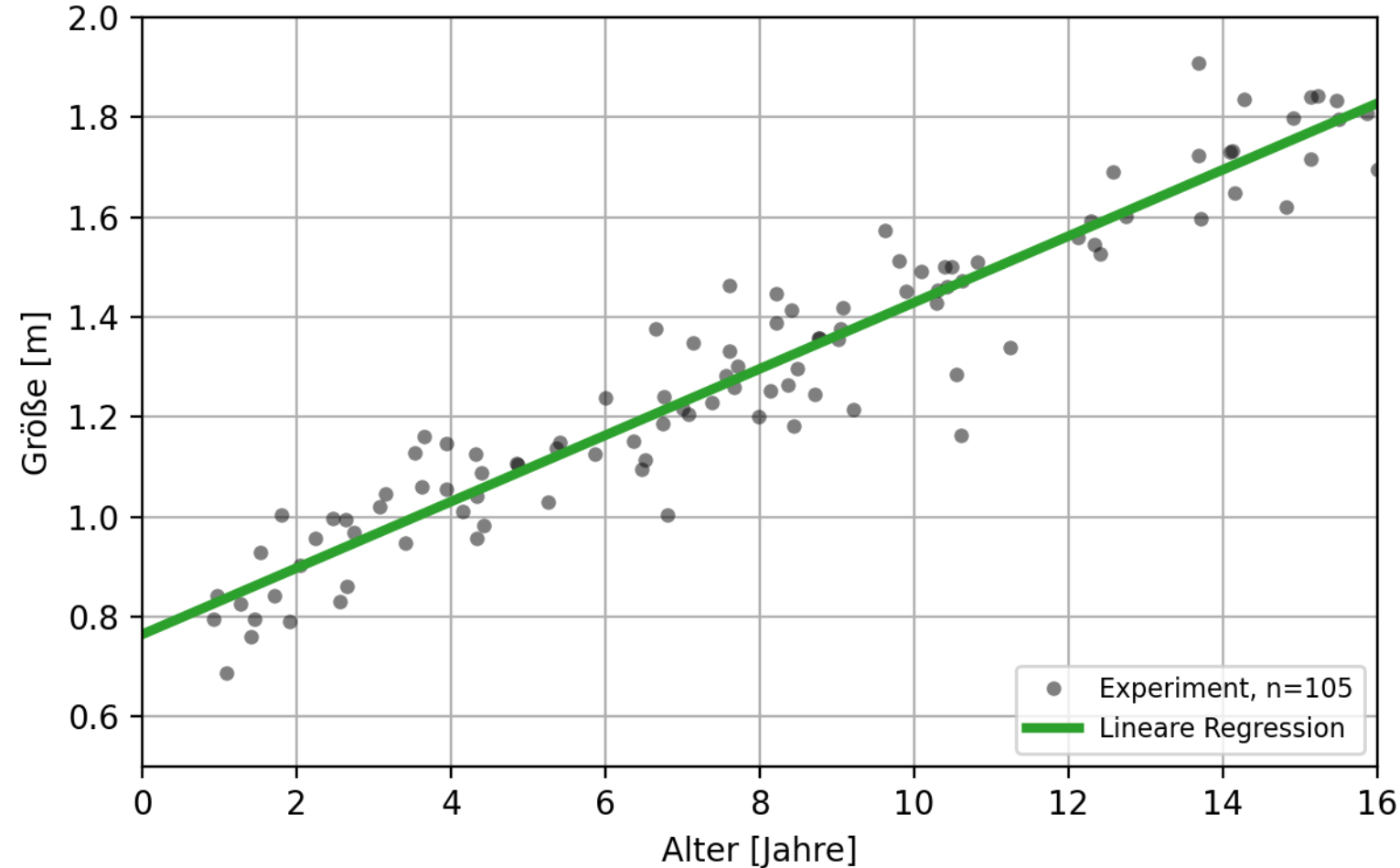
Lineare Regression

Wachstum von Kindern



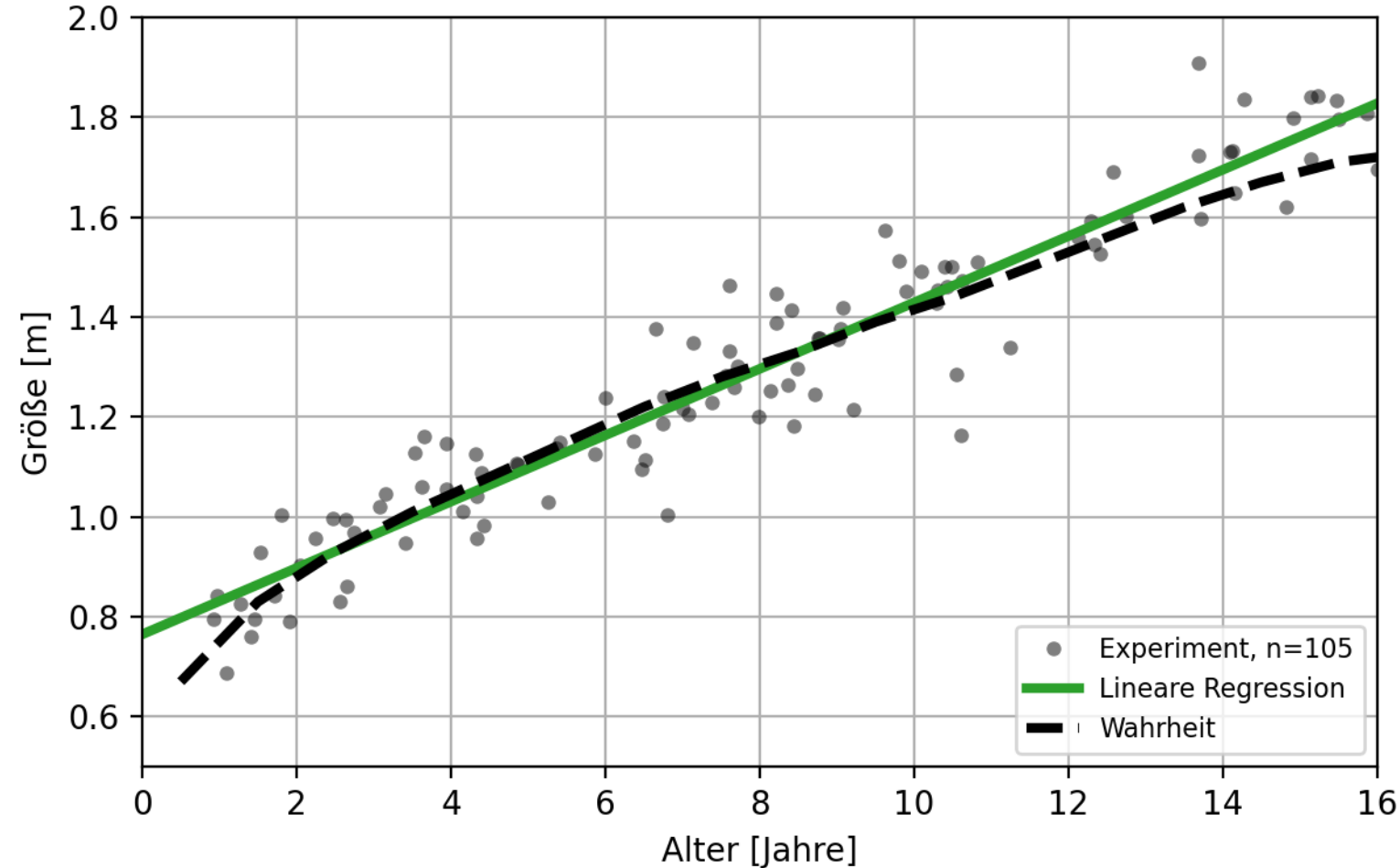
Lineare Regression

Wachstum von Kindern



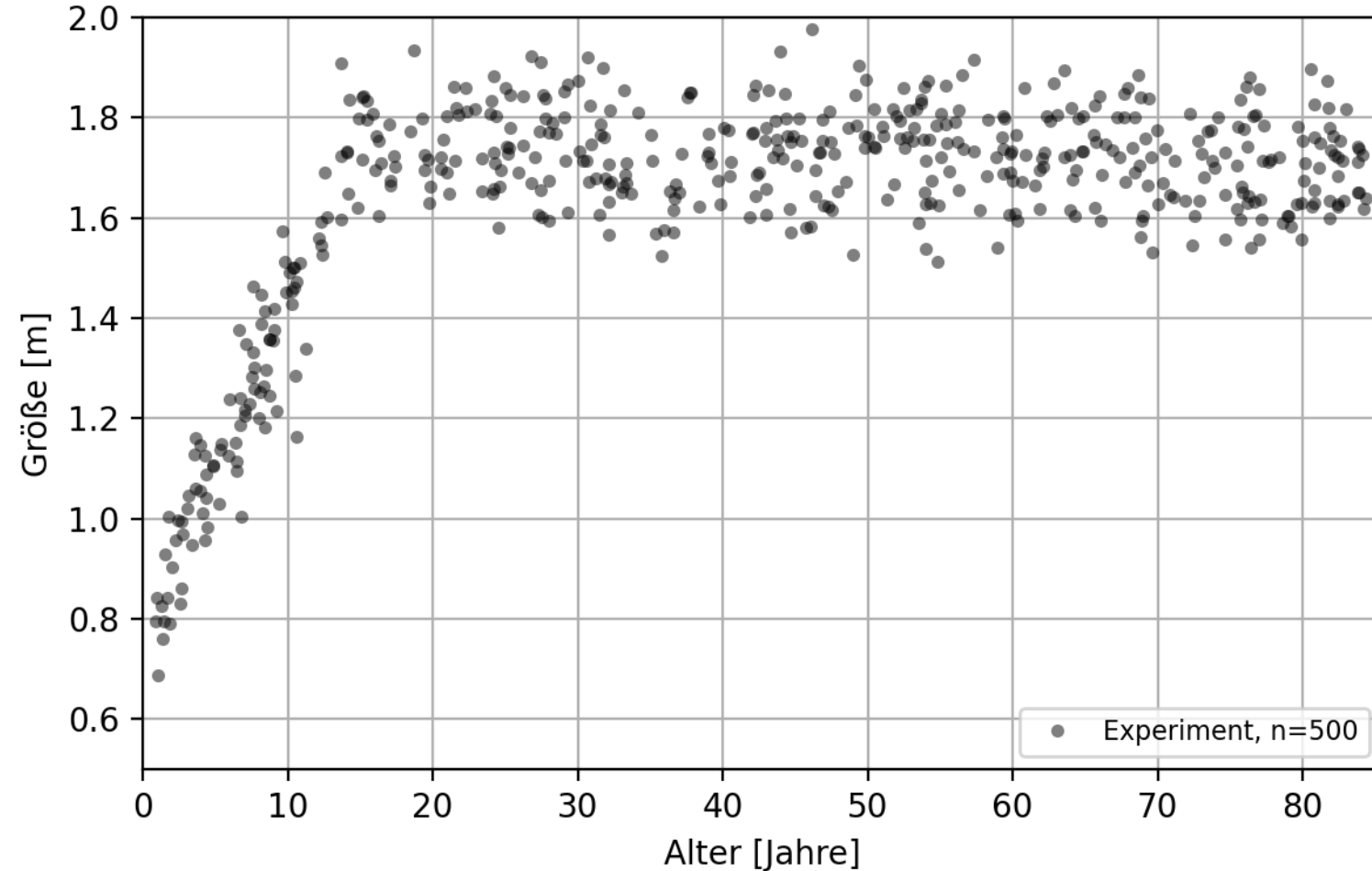
Lineare Regression

Wachstum von Kindern



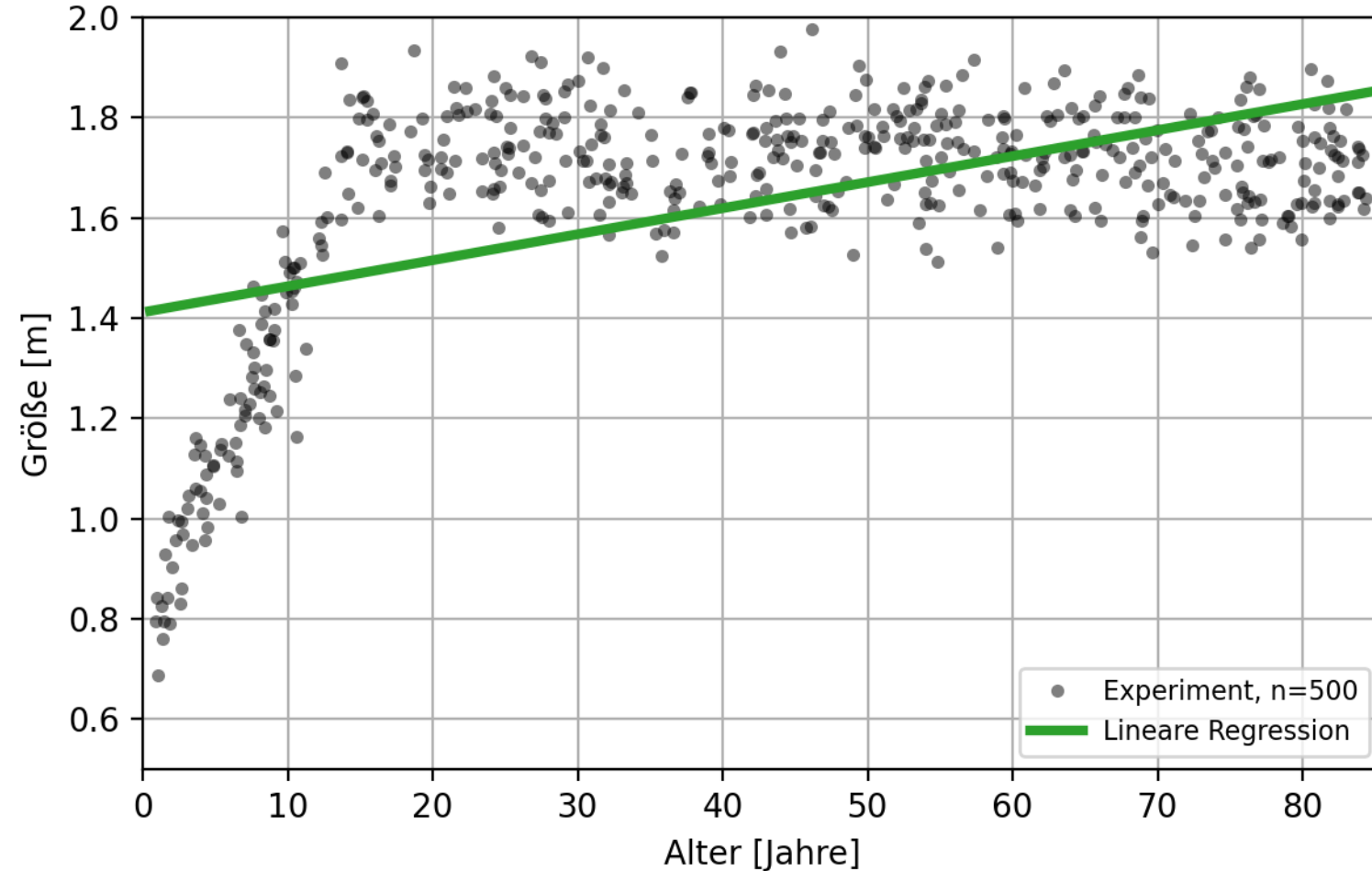
Quelle:
**Gesundheitsberichterstattung
der Bundes**
www.gbe-bund.de

Überblick Regressoren



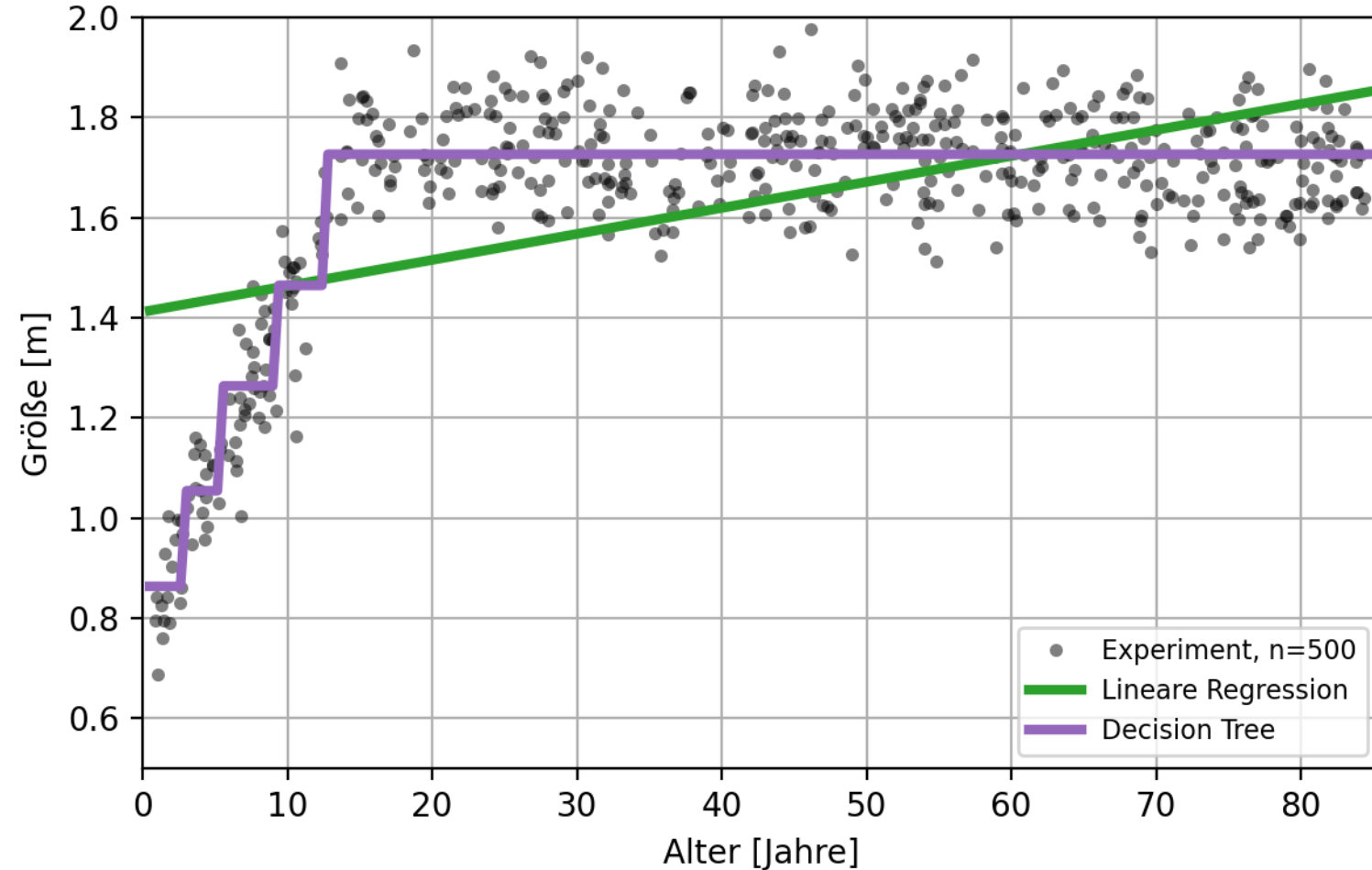
Quelle:
**Gesundheitsberichterstattung
der Bundes**
www.gbe-bund.de

Überblick Regressoren



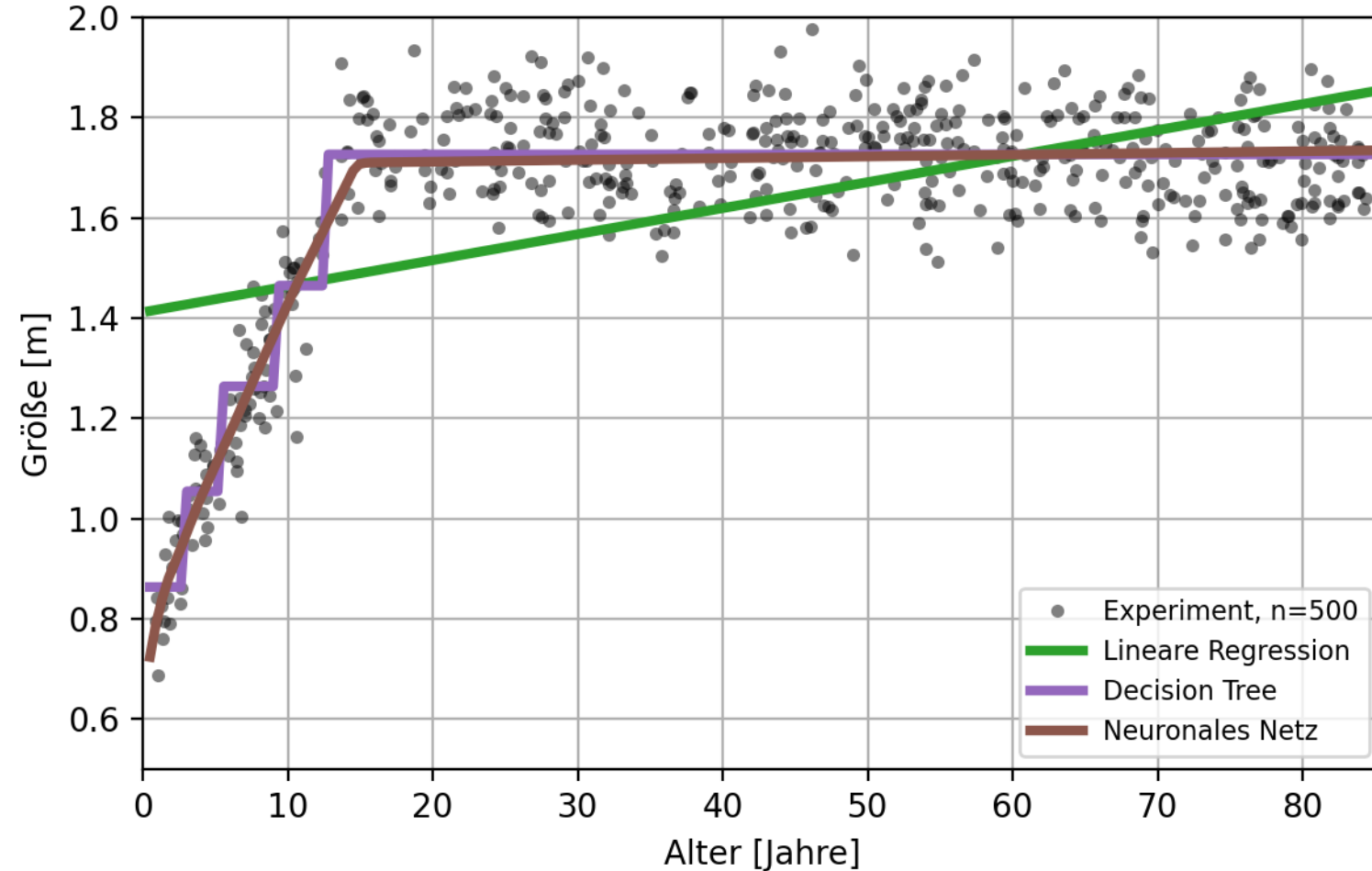
Quelle:
**Gesundheitsberichterstattung
der Bundes**
www.gbe-bund.de

Überblick Regressoren



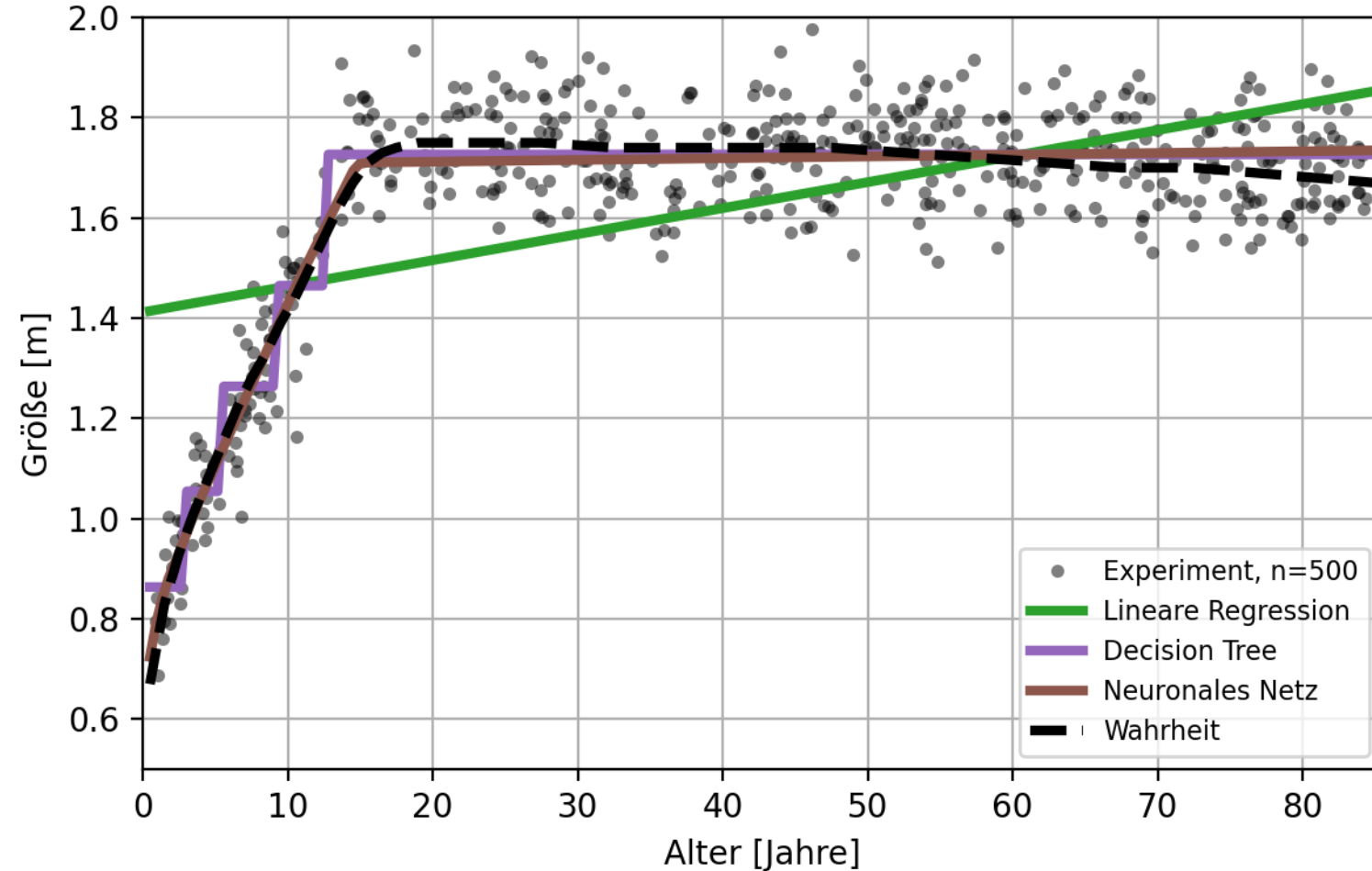
Quelle:
**Gesundheitsberichterstattung
der Bundes**
www.gbe-bund.de

Überblick Regressoren



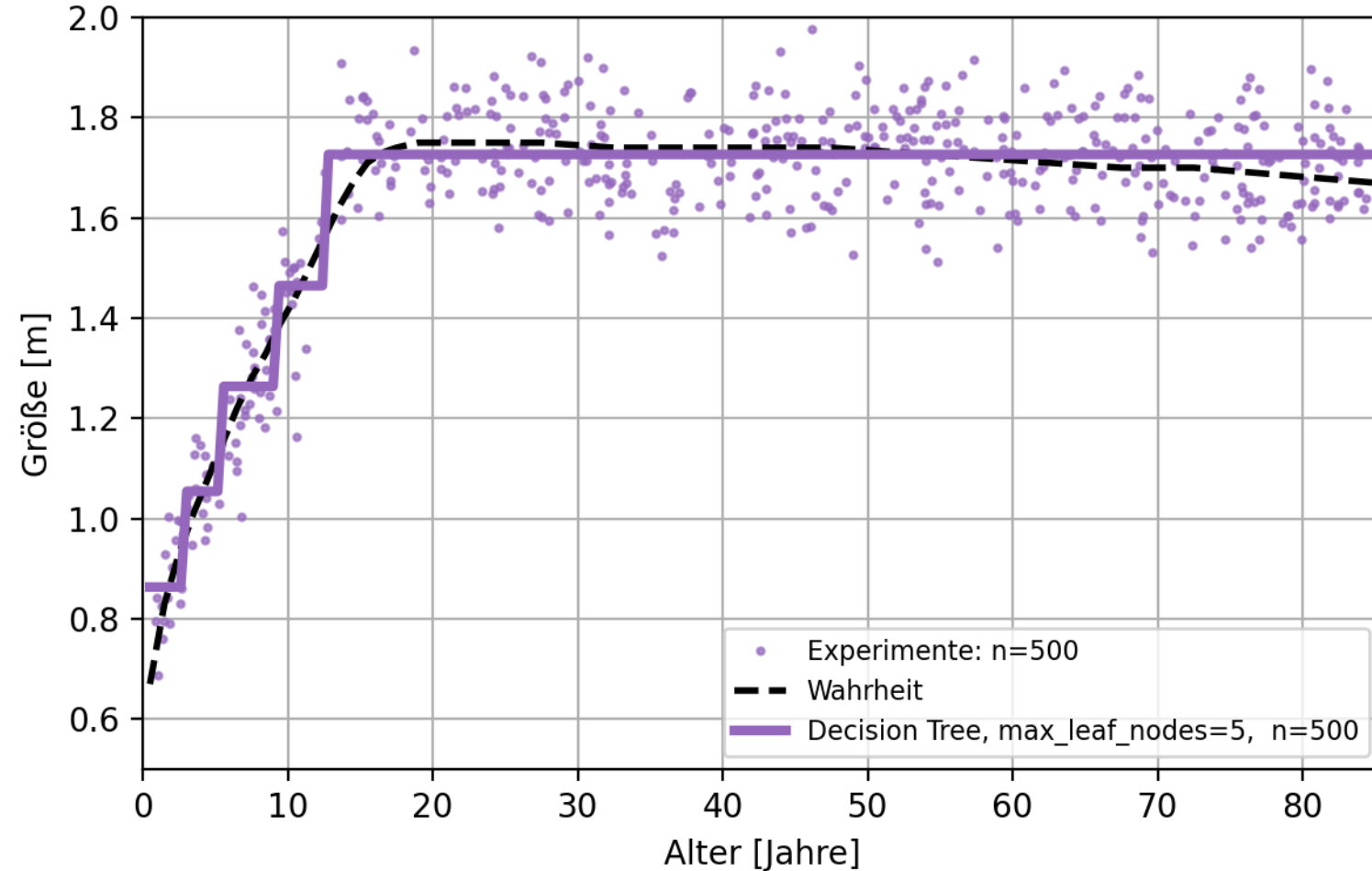
Quelle:
**Gesundheitsberichterstattung
der Bundes**
www.gbe-bund.de

Überblick Regressoren



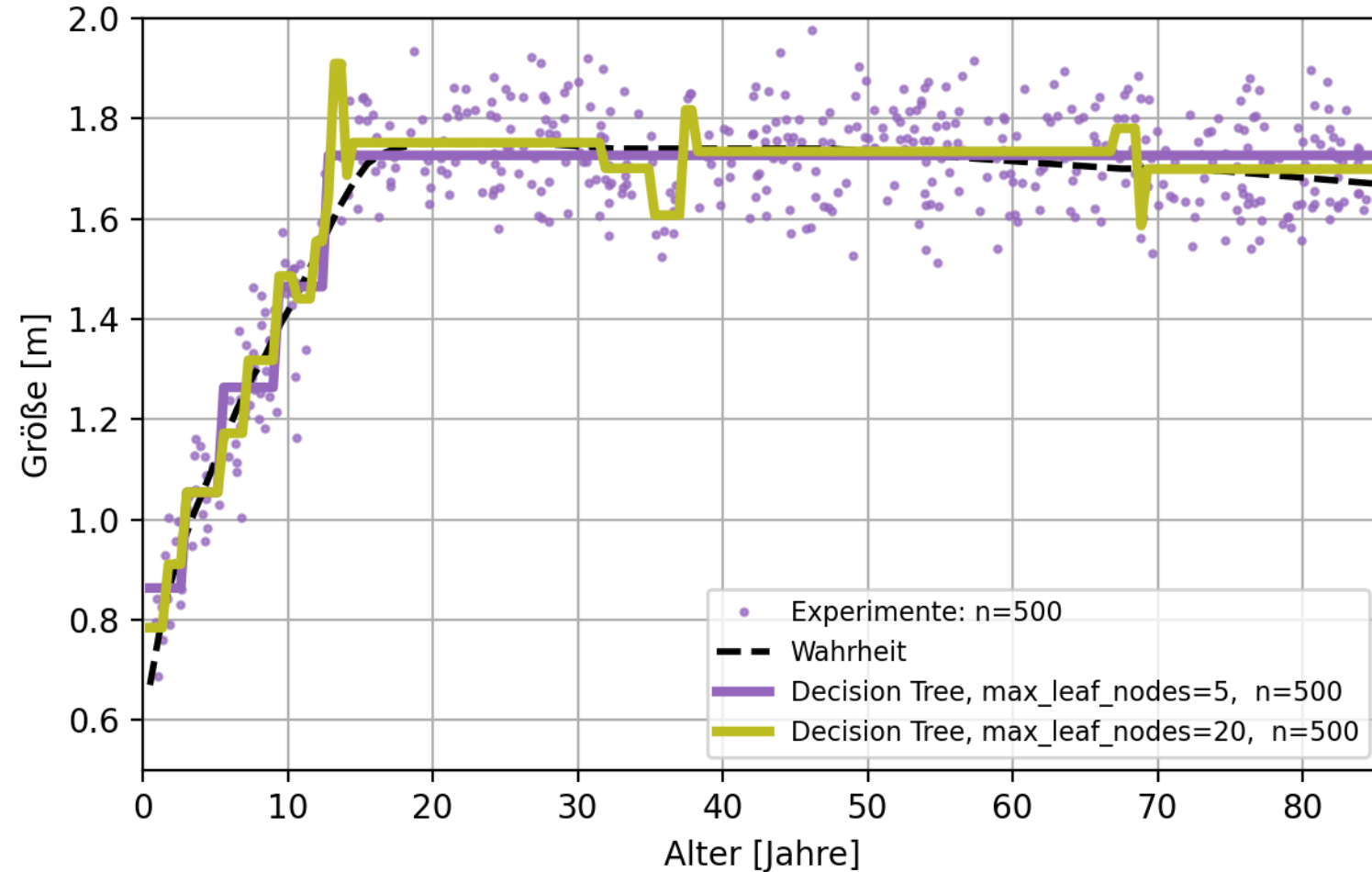
Quelle:
**Gesundheitsberichterstattung
der Bundes**
www.gbe-bund.de

Decision Tree



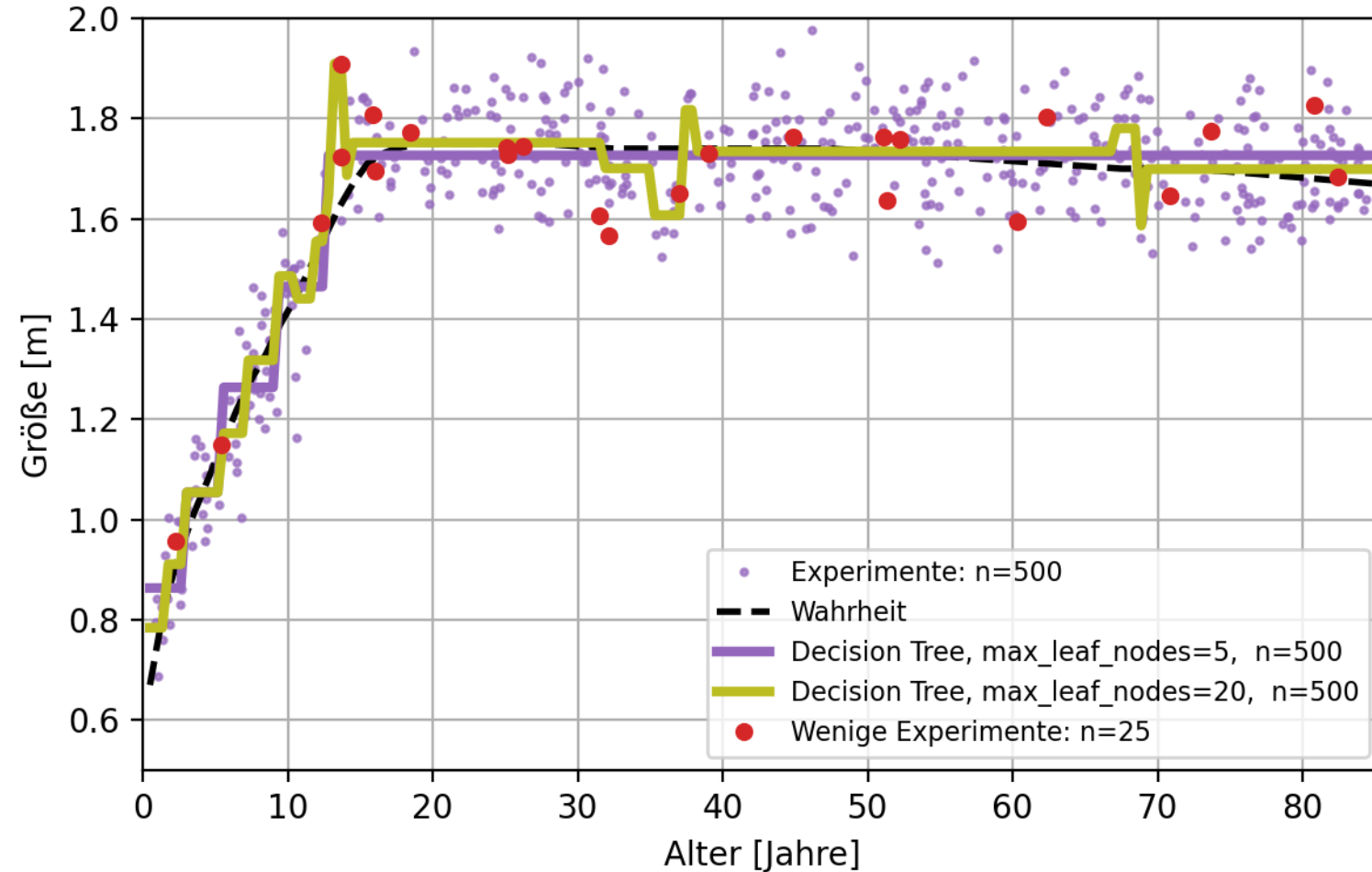
Quelle:
**Gesundheitsberichterstattung
der Bundes**
www.gbe-bund.de

Decision Tree



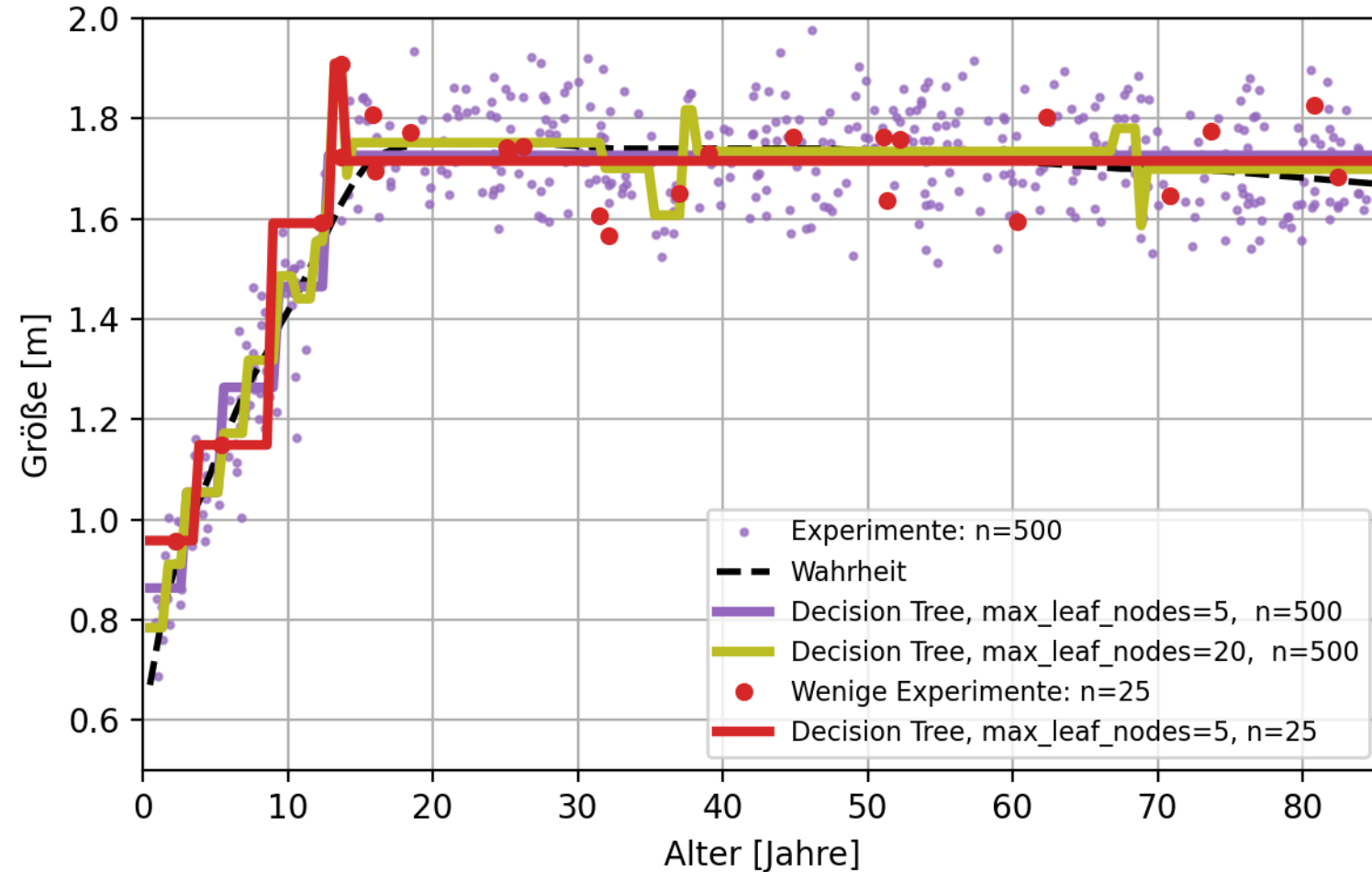
Quelle:
**Gesundheitsberichterstattung
der Bundes**
www.gbe-bund.de

Decision Tree



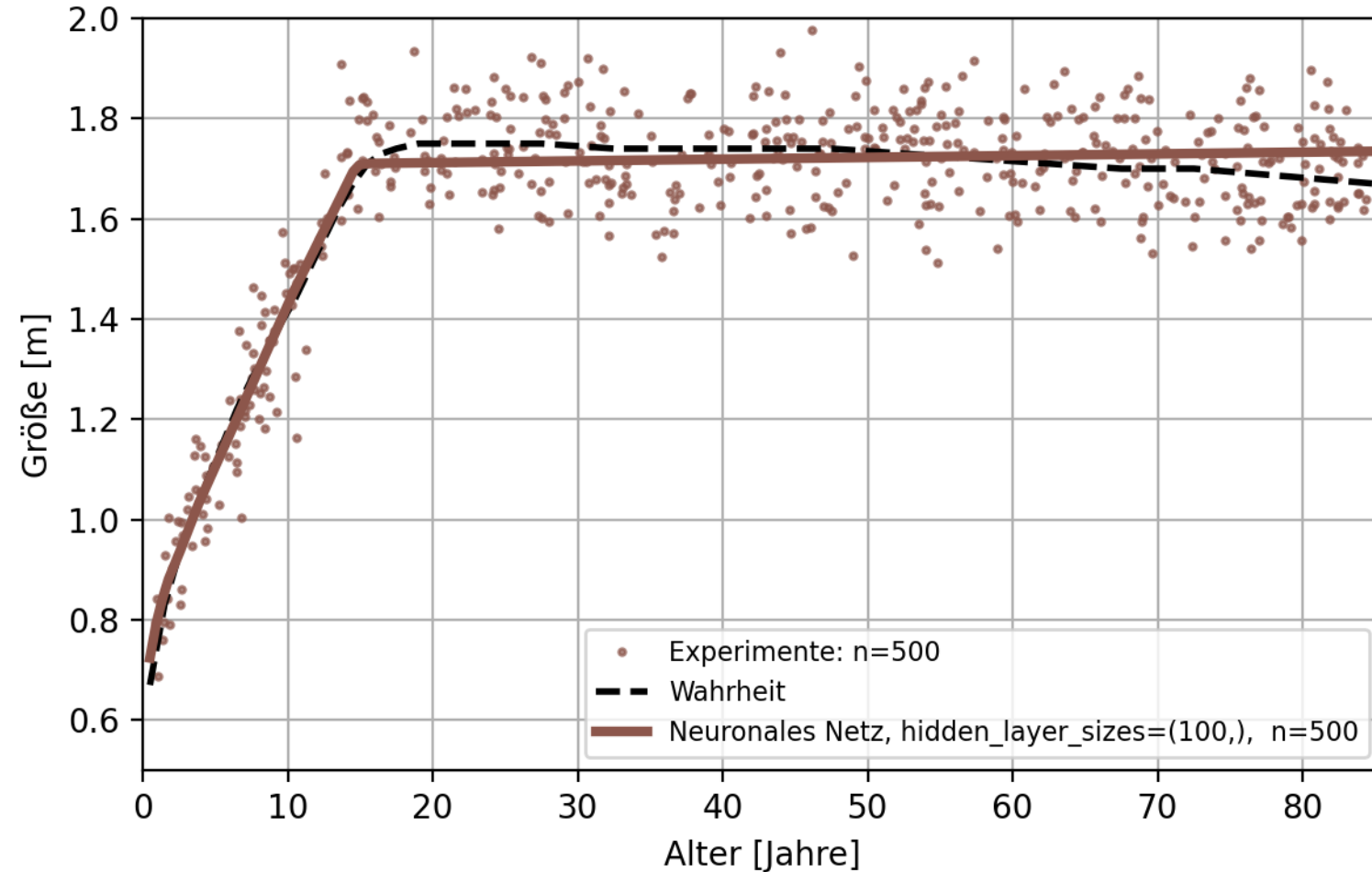
Quelle:
**Gesundheitsberichterstattung
der Bundes**
www.gbe-bund.de

Decision Tree



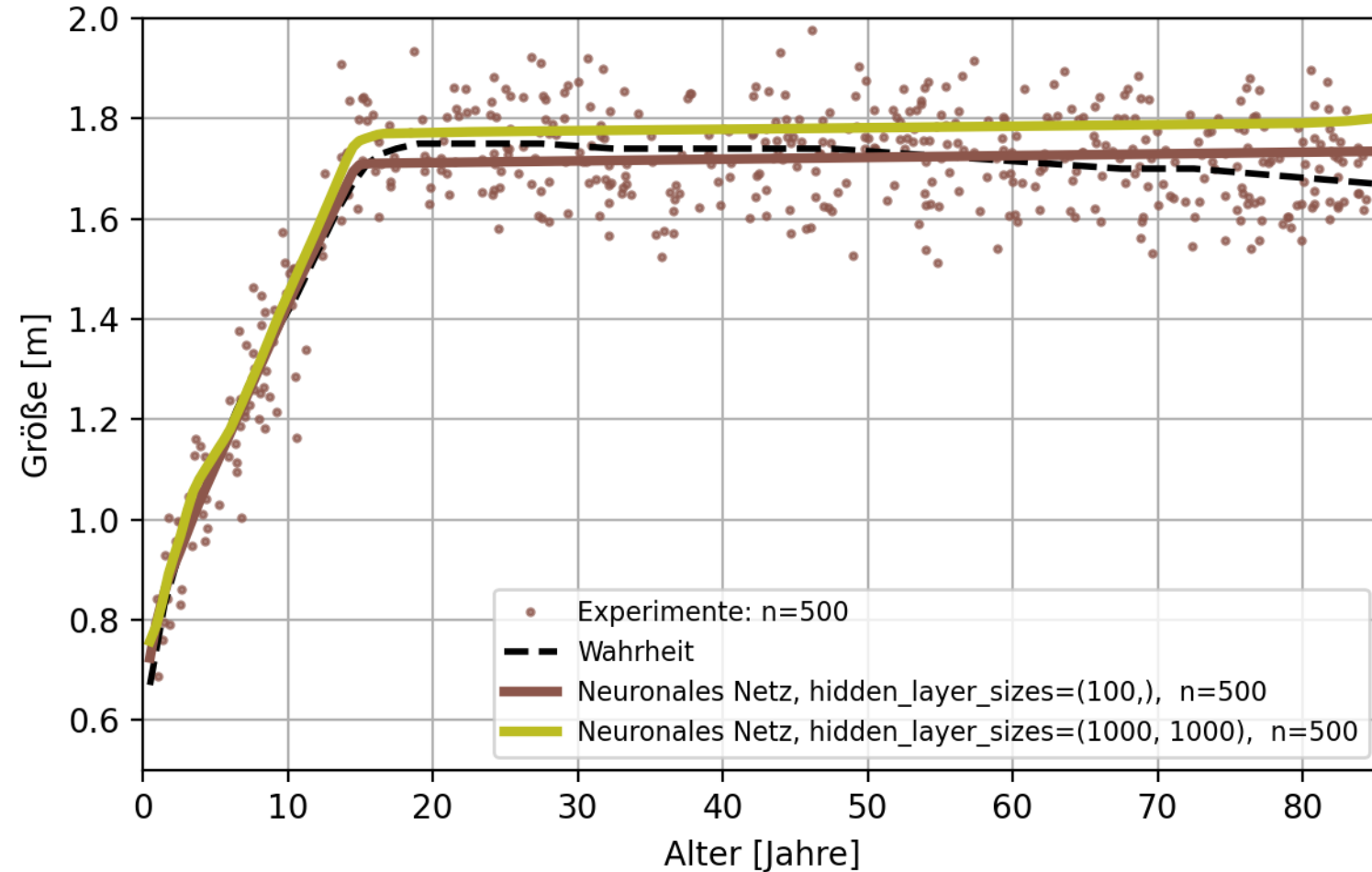
Quelle:
**Gesundheitsberichterstattung
der Bundes**
www.gbe-bund.de

Neuronales Netz



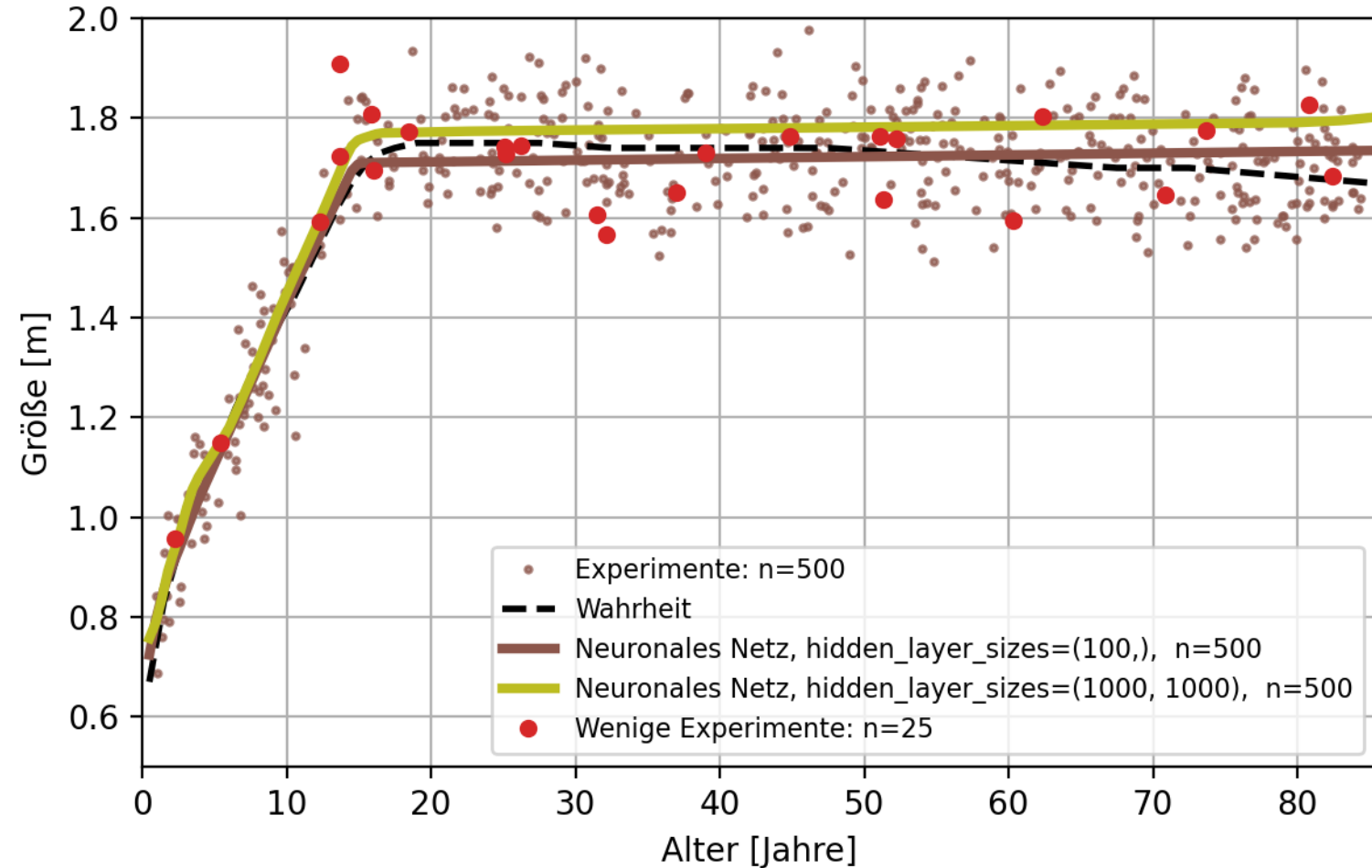
Quelle:
**Gesundheitsberichterstattung
der Bundes**
www.gbe-bund.de

Neuronales Netz



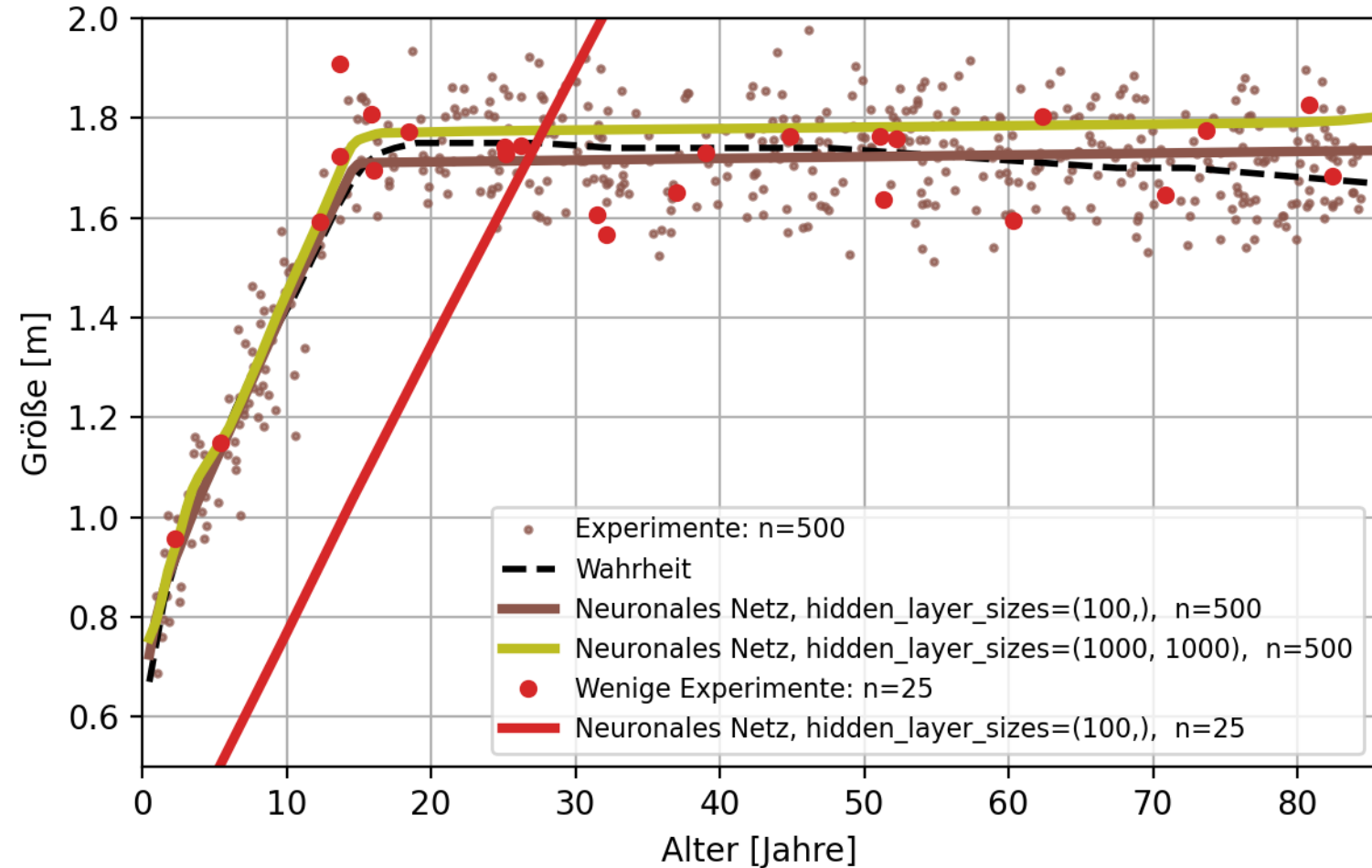
Quelle:
**Gesundheitsberichterstattung
der Bundes**
www.gbe-bund.de

Neuronales Netz



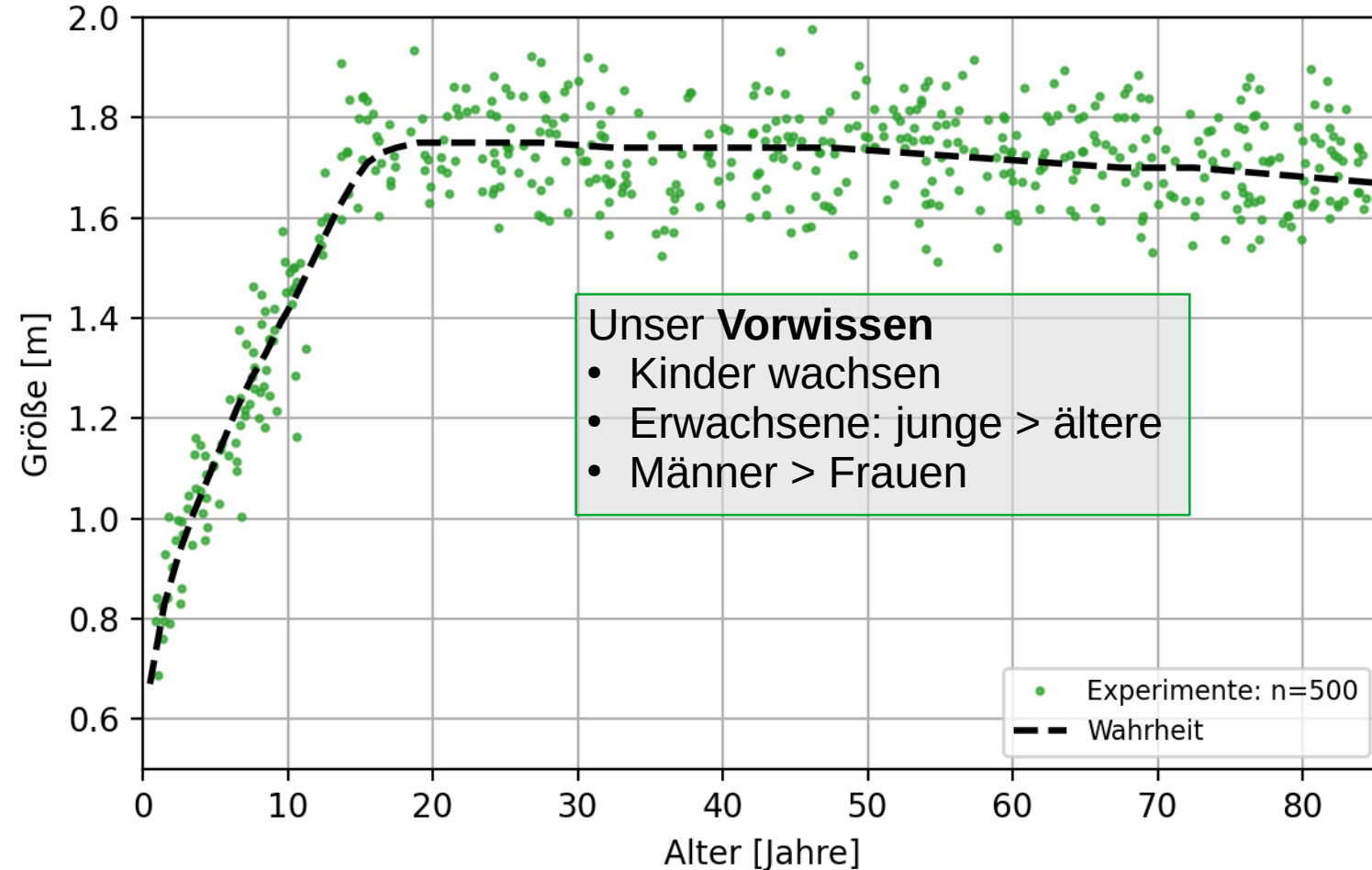
Quelle:
**Gesundheitsberichterstattung
der Bundes**
www.gbe-bund.de

Neuronales Netz



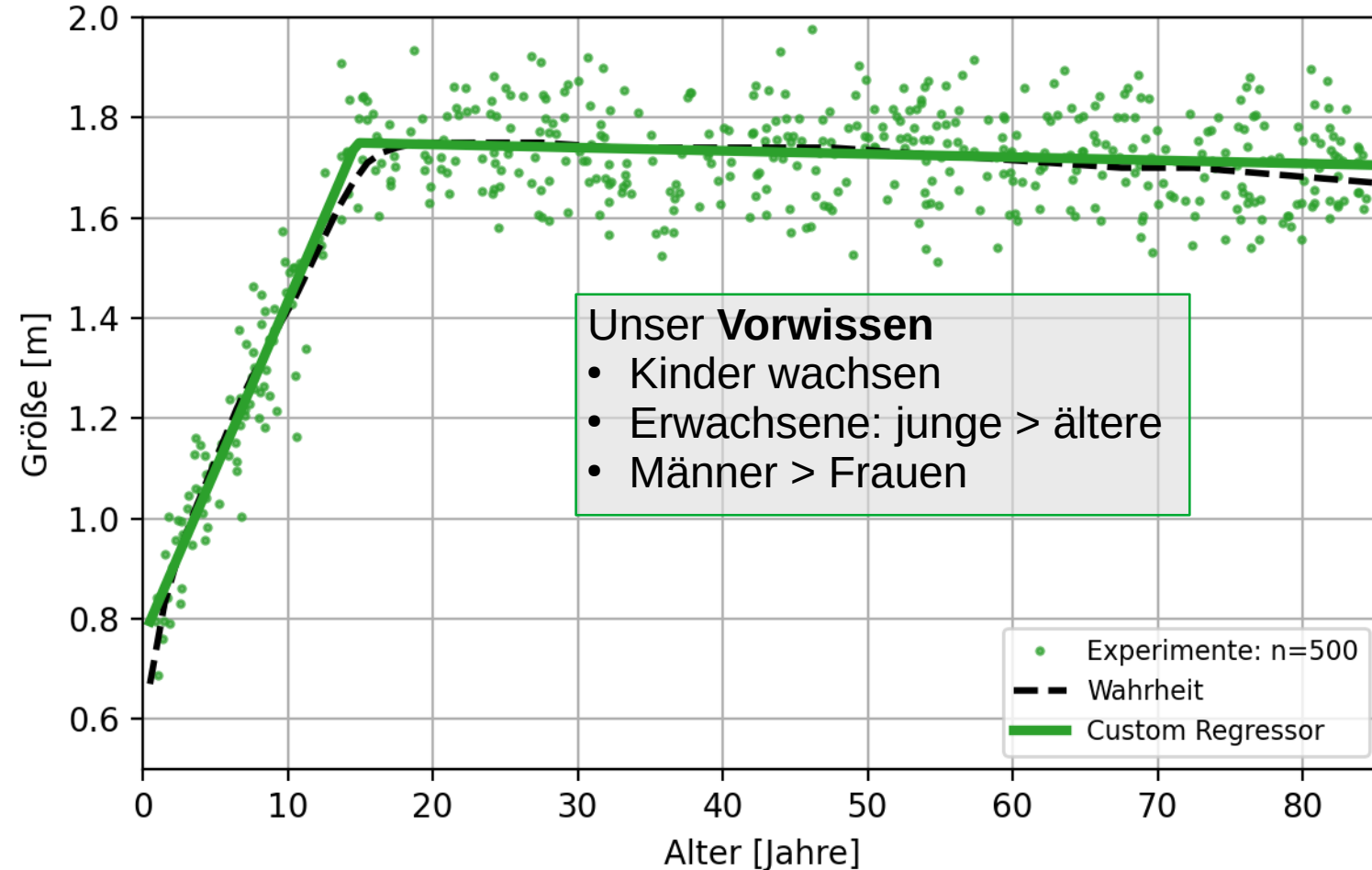
Quelle:
**Gesundheitsberichterstattung
der Bundes**
www.gbe-bund.de

Individuallösung



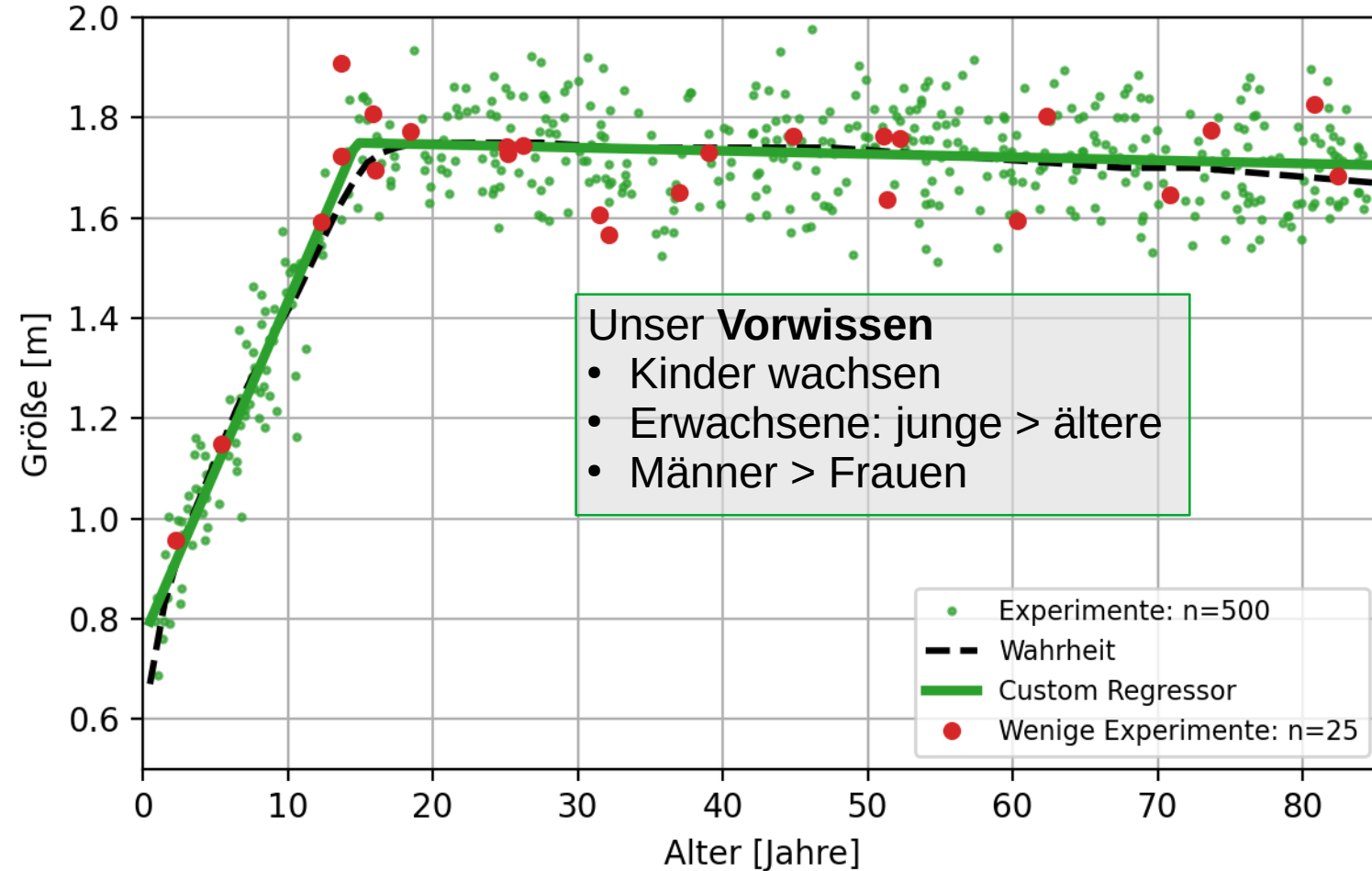
Quelle:
**Gesundheitsberichterstattung
der Bundes**
www.gbe-bund.de

Individuallösung



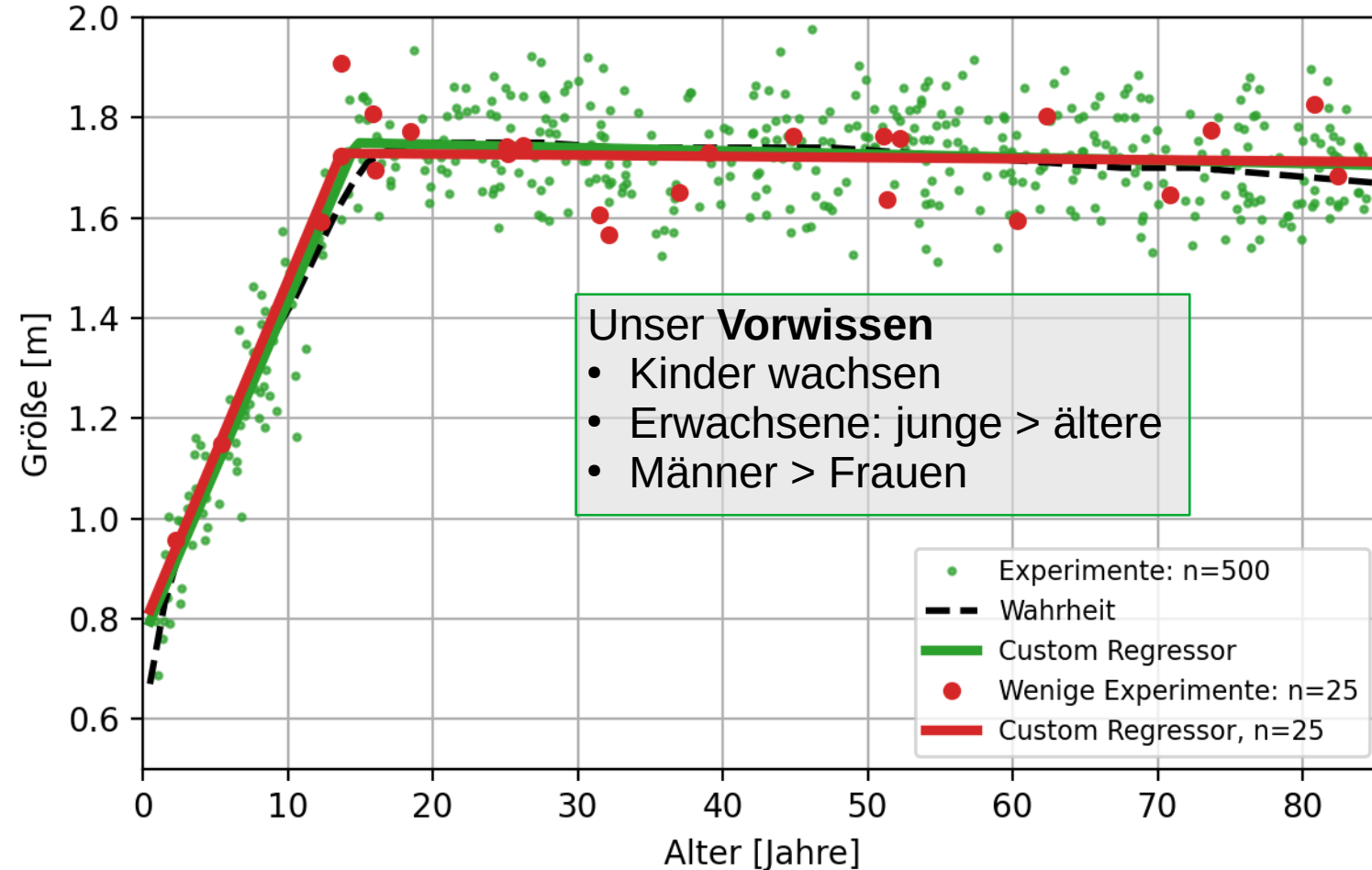
Quelle:
**Gesundheitsberichterstattung
der Bundes**
www.gbe-bund.de

Individuallösung



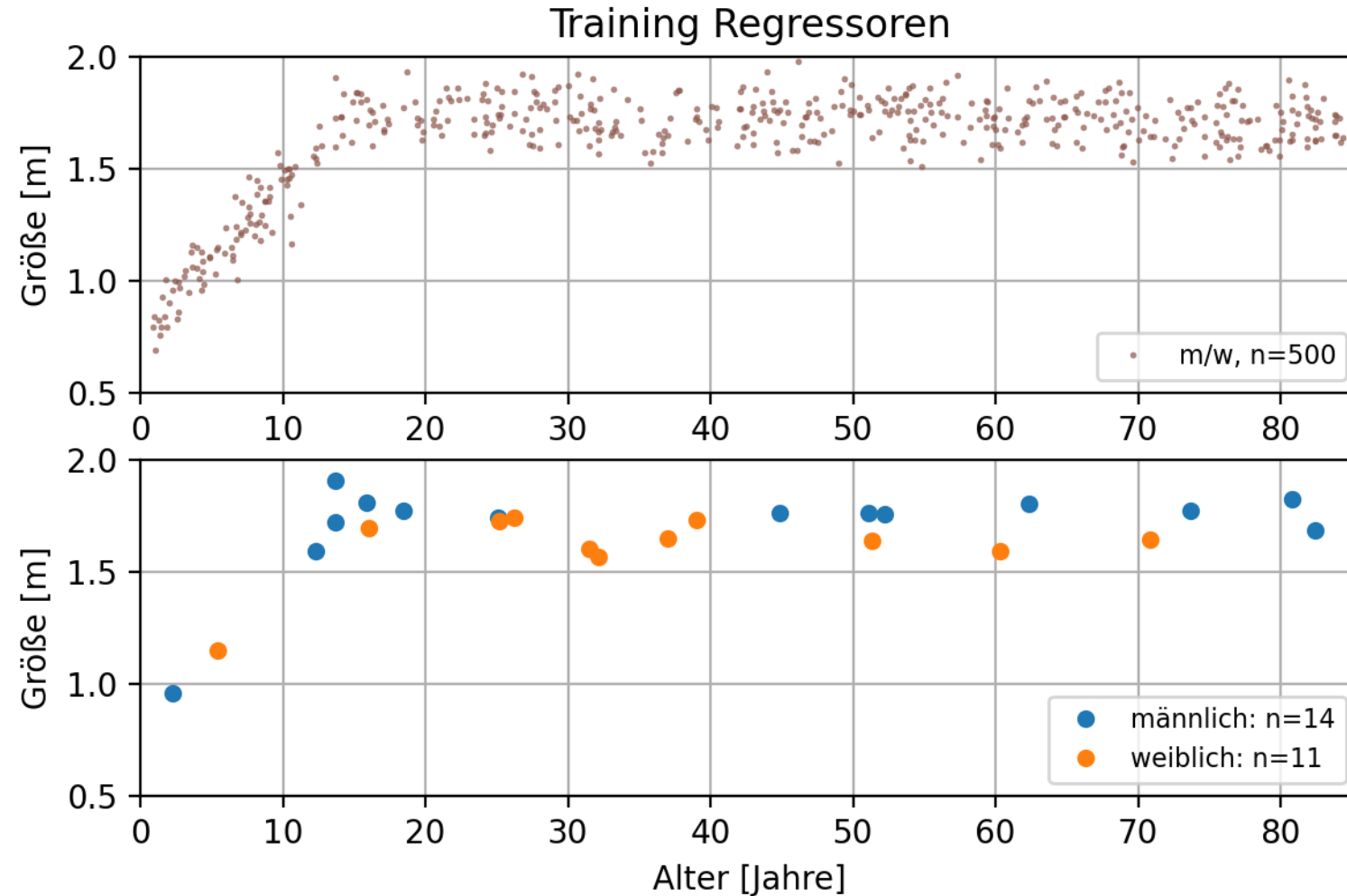
Quelle:
**Gesundheitsberichterstattung
der Bundes**
www.gbe-bund.de

Individuallösung



Quelle:
**Gesundheitsberichterstattung
der Bundes**
www.gbe-bund.de

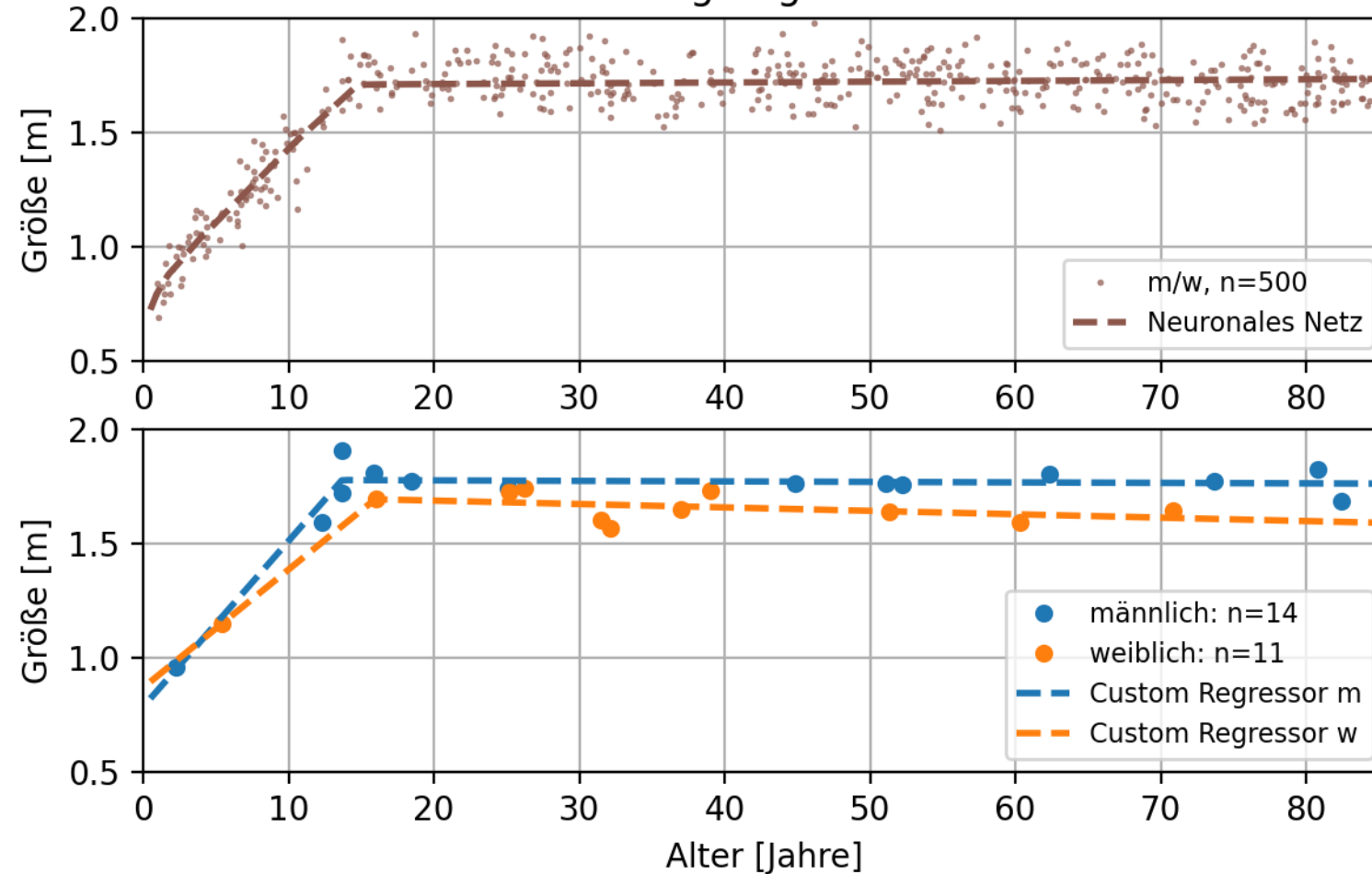
Benchmark



Quelle:
**Gesundheitsberichterstattung
der Bundes**
www.gbe-bund.de

Benchmark

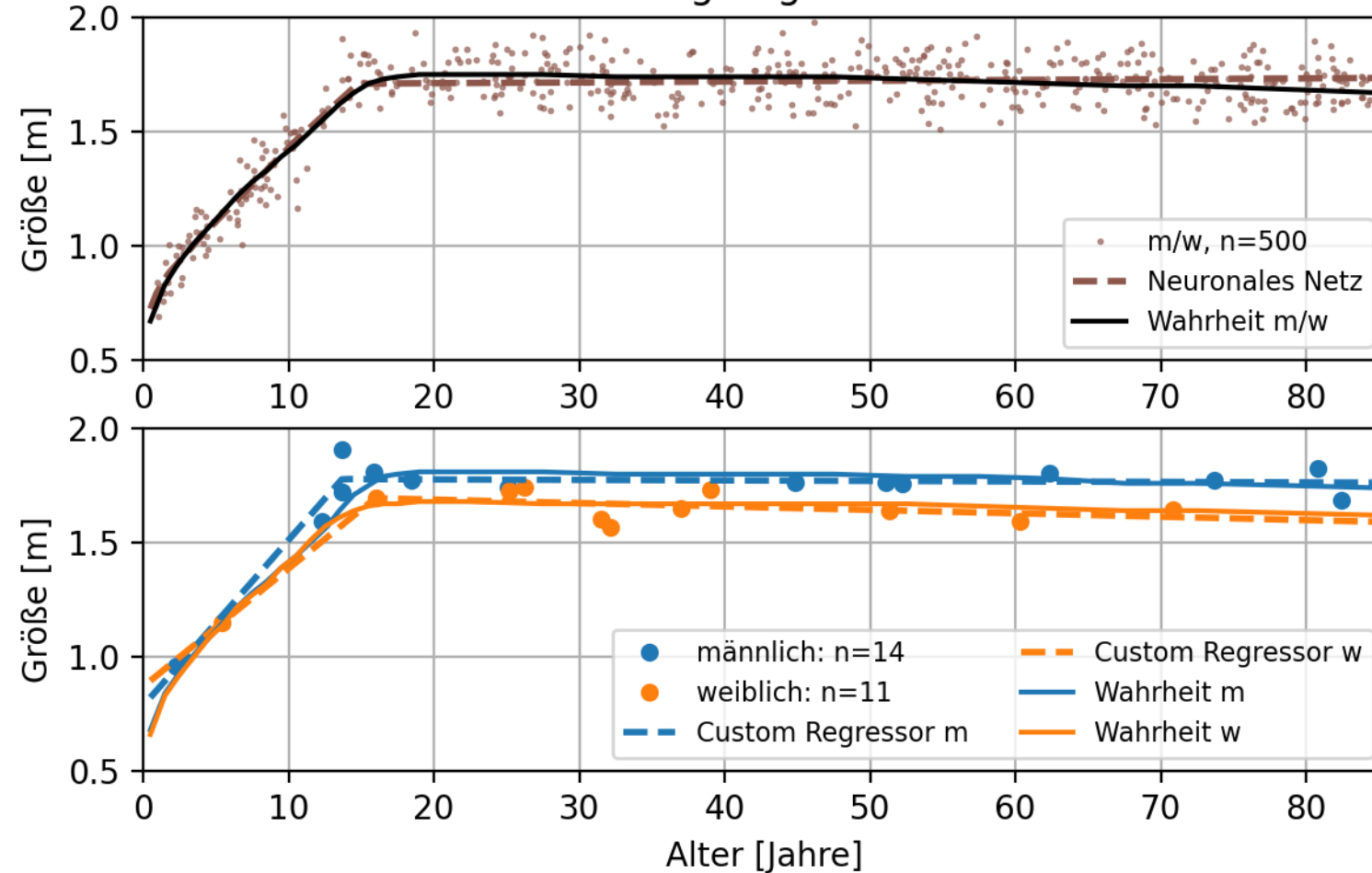
Training Regressoren



Quelle:
**Gesundheitsberichterstattung
der Bundes**
www.gbe-bund.de

Benchmark

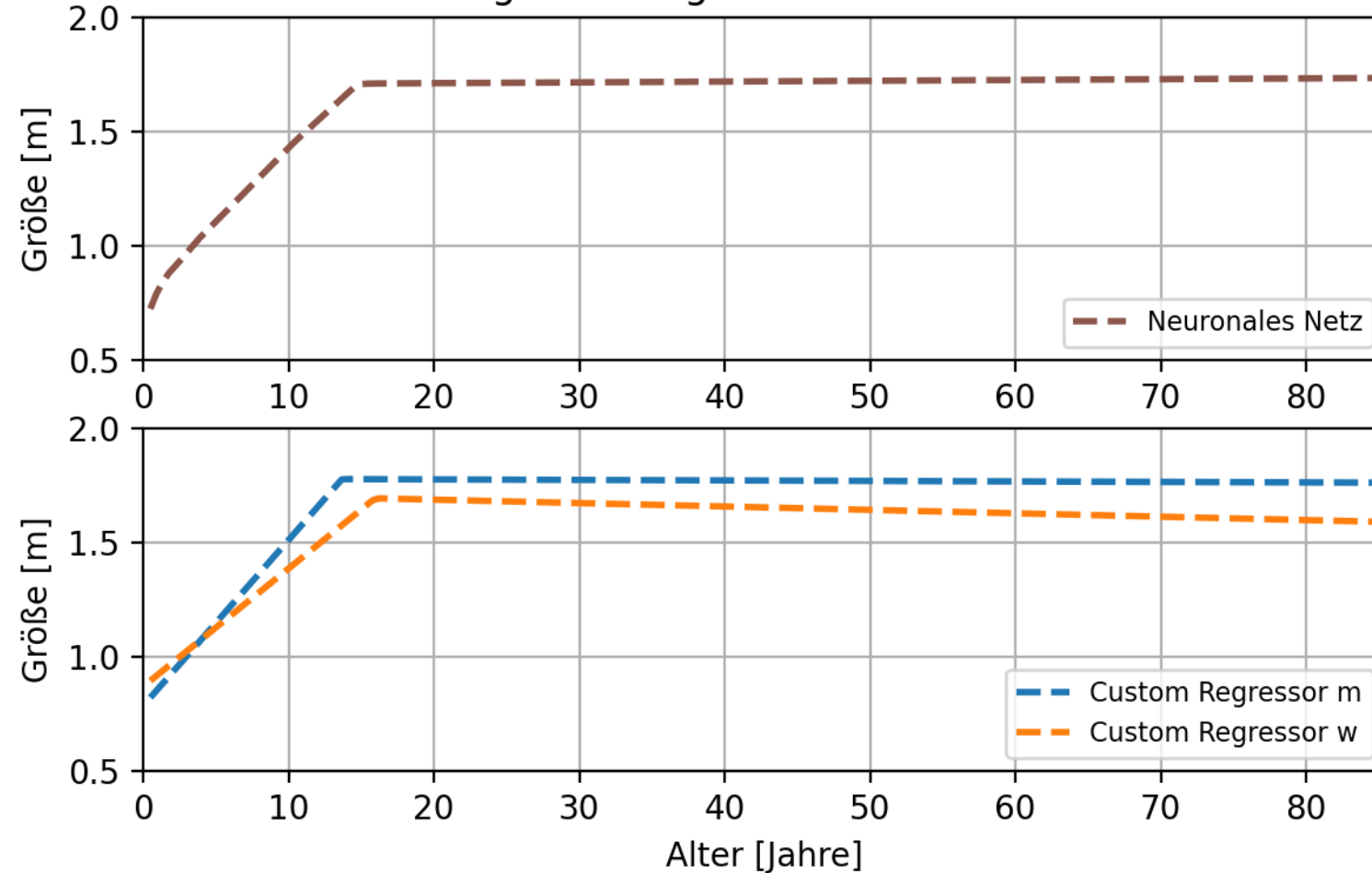
Training Regressoren



Quelle:
**Gesundheitsberichterstattung
der Bundes**
www.gbe-bund.de

Benchmark

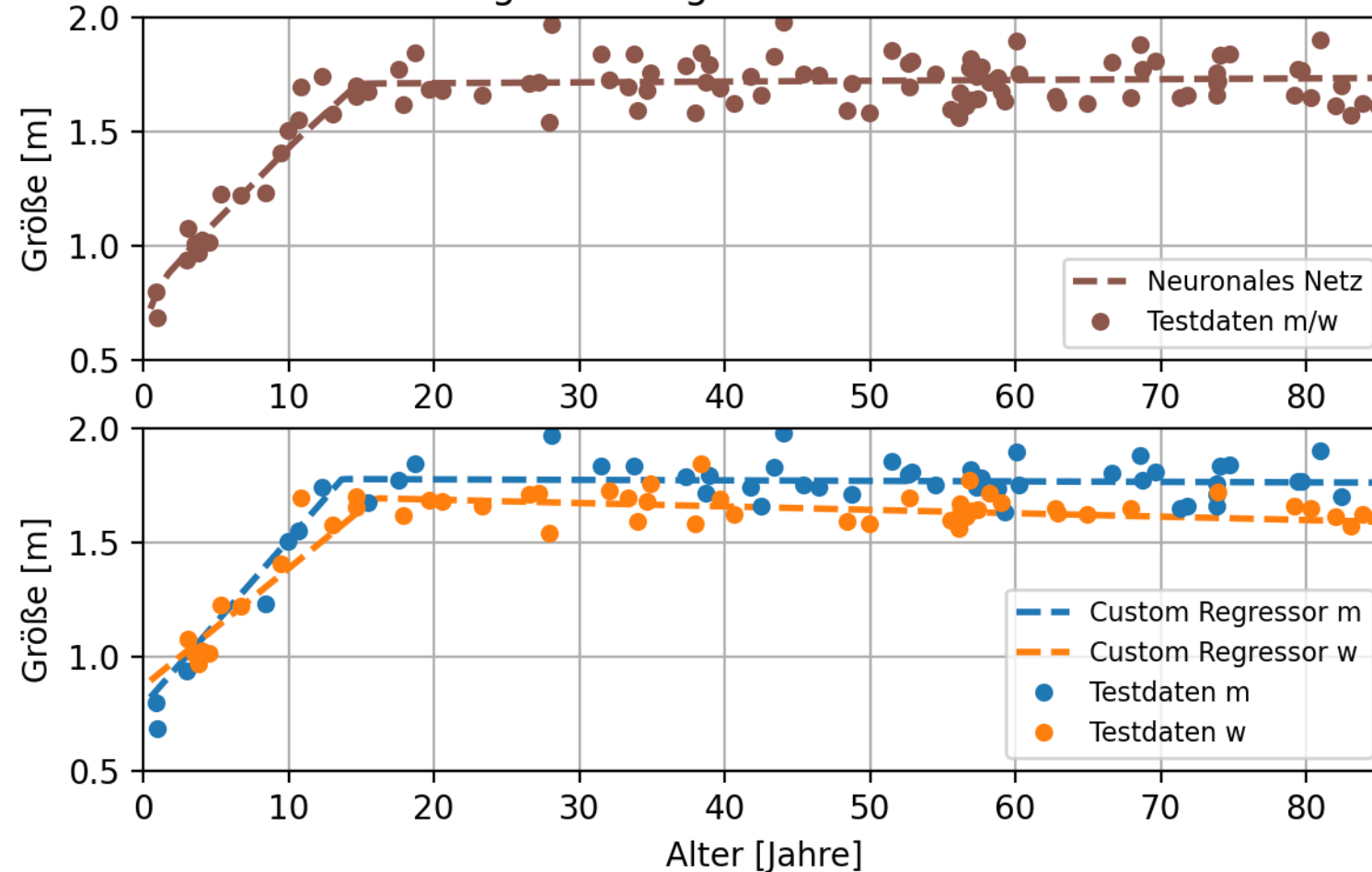
Vergleich: Regressor vs. Testdaten



Data Science
Lösungen
via Wettbewerb:
www.kaggle.com

Benchmark

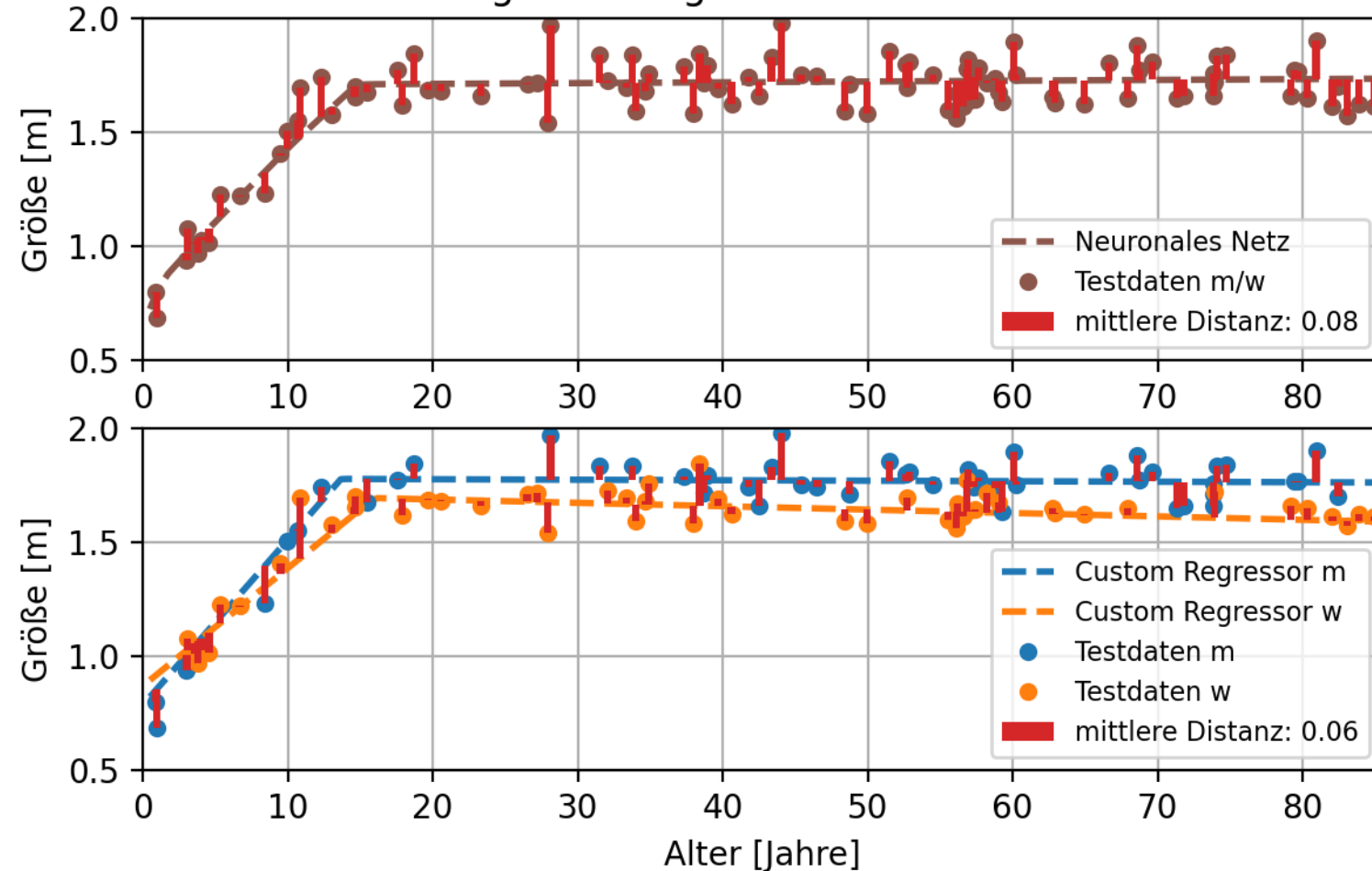
Vergleich: Regressor vs. Testdaten



Data Science
Lösungen
via Wettbewerb:
www.kaggle.com

Benchmark

Vergleich: Regressor vs. Testdaten



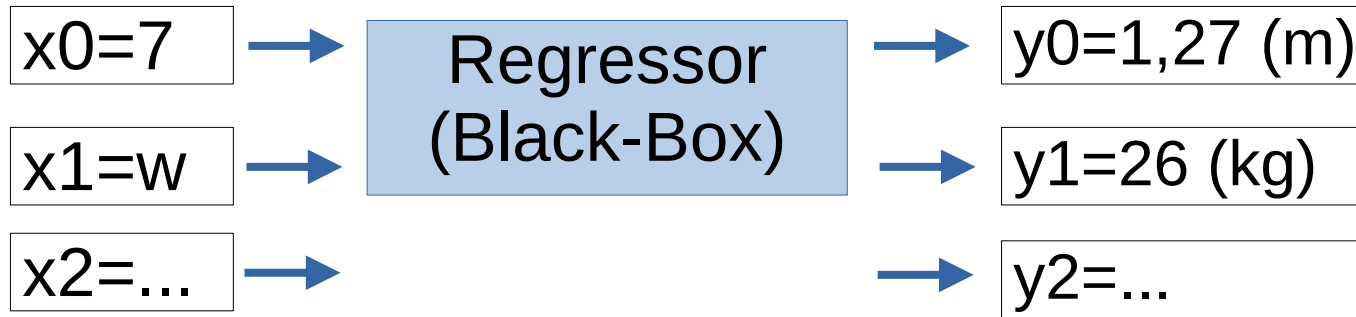
Data Science
Lösungen
via Wettbewerb:
www.kaggle.com

Zwischenfazit

Neuronale Netze sind nicht die Universallösung

	komplexes Machine Learning Model	Lösung basierend auf Systemverständnis
Entwicklungsaufwand	niedrig	hoch
Bedarf an Trainingsdaten	hoch	niedrig
Wiederverwertbarkeit	niedrig	hoch
Vorhersageverlässlichkeit	allgemein: gut selten: Totalversagen	sehr gut
Unschlagbar bei	Bild, Ton, Video, Ästhetik	Extreme Verlässlichkeit

Verallgemeinerung



Beispiele:

- Aus Gewicht (x_0), aktuellem Ort (x_1, x_2, x_3), Geschwindigkeit (x_4, x_5, x_6) und Windstärke (x_7, x_8, x_9) von brennenden Spänen
... sage Auftreffpunkt (y_0, y_1, y_2) und Streuung (y_3, y_4, y_5) vorher.
- Aus Temperatur (x_0), Lackart (x_1), Untergrundmaterial (x_2) und Zeit (x_3)
... sage Trocknungsgrad (y_0) vorher.
- Aus den Ergebnissen der Produktprüfung ($x_0, \dots$)
... sage Lebensdauer (y_0) vorher.

NN @ Handschrifterkennung

5 0 4 1 9 2

- NIST Data-Set „handwritten digits“
- 250 Autoren aus US Census Bureau
- Training: **60 000 Ziffern**
- Test-Daten: 10 000 Ziffern

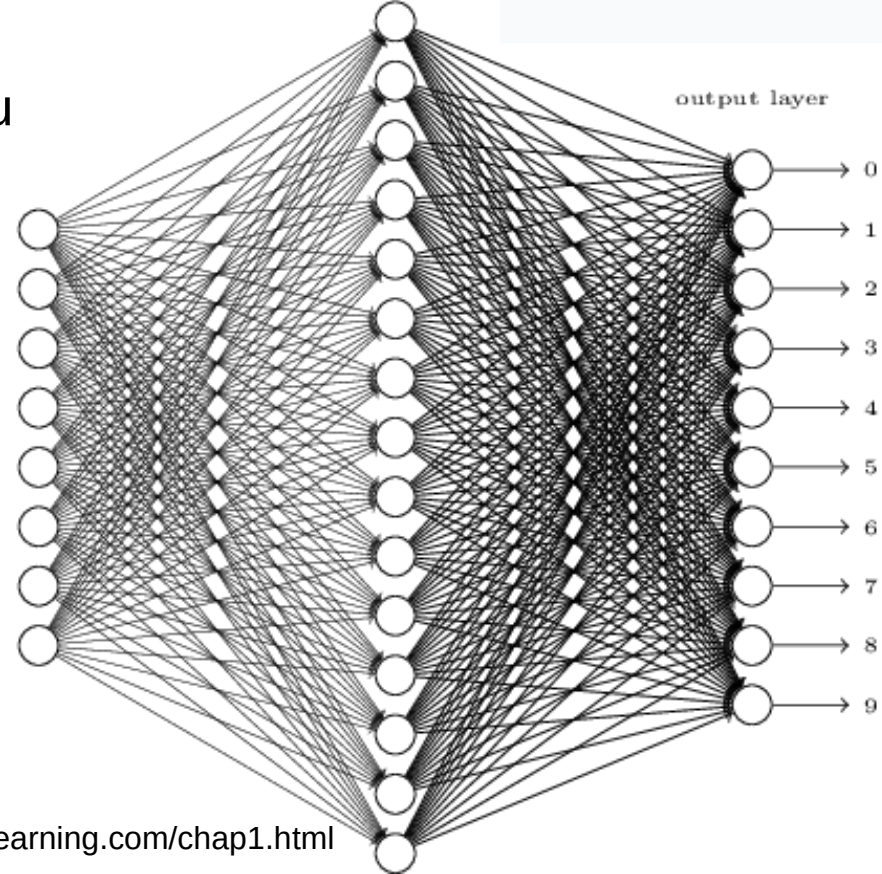
28 px
28 px



input layer
(784 neurons)

hidden layer
(n = 15 neurons)

output layer

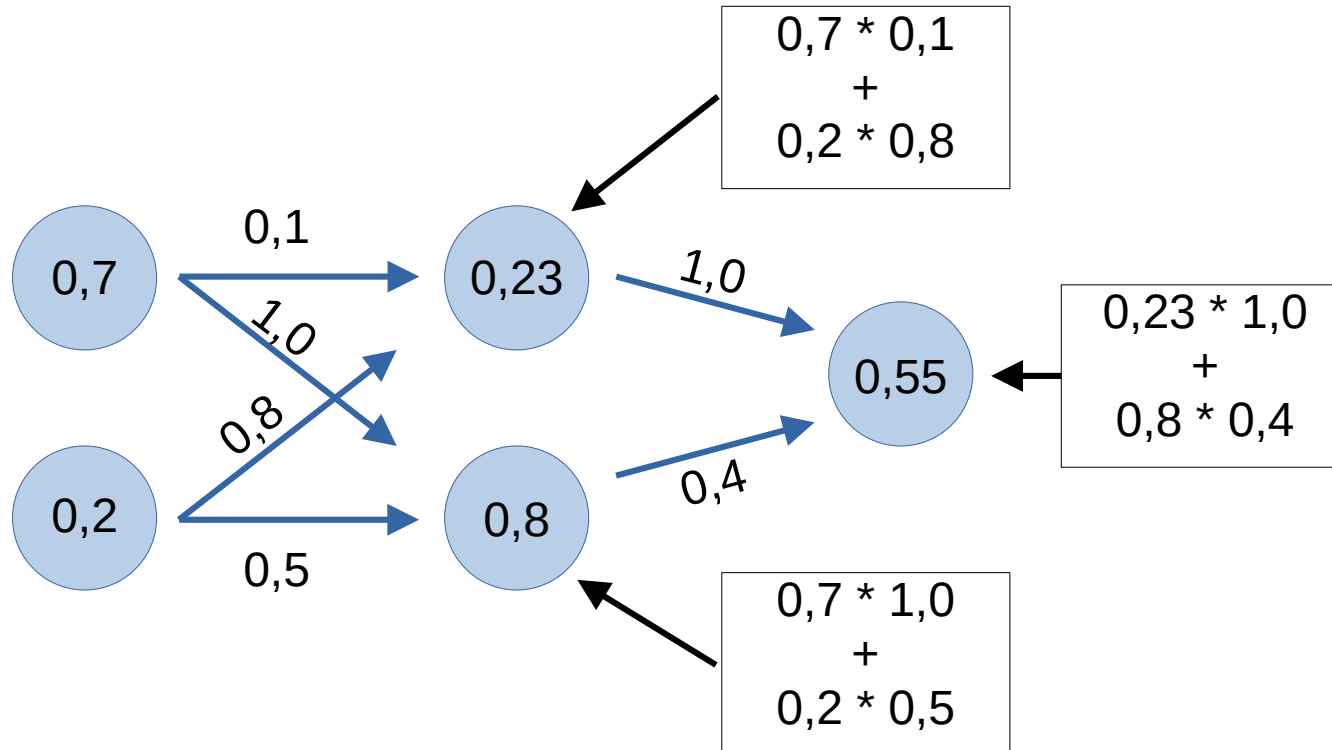


- Genauigkeit: 99,79%
(<https://proceedings.mlr.press/v28/wan13.pdf>)
- Typische Genauigkeit: 20% - 50%

<http://neuralnetworksanddeeplearning.com/chap1.html>

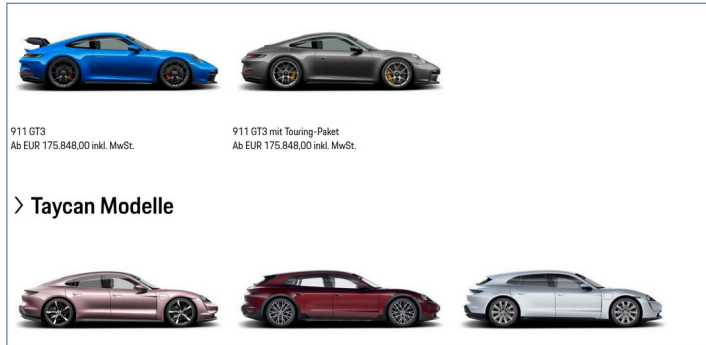
Neuronales Netz

Funktionsweise (vereinfacht):

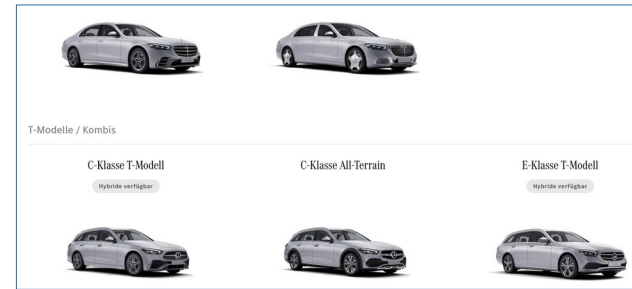


Aufgabe: Finde Automarke

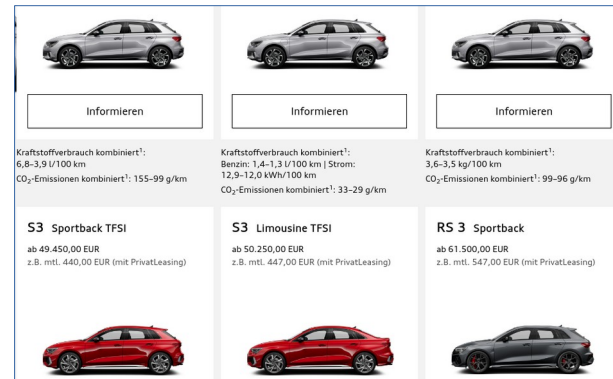
<https://www.porsche.com/germany/models/>



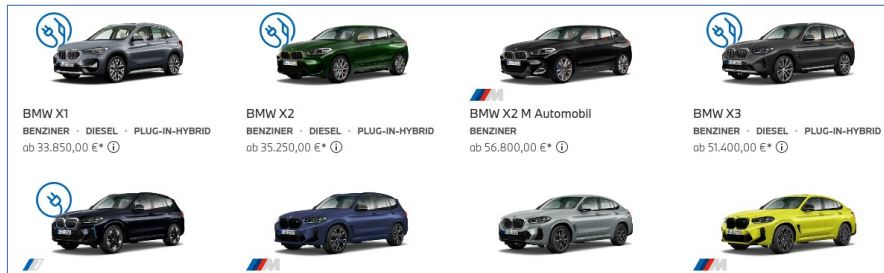
<https://www.mercedes-benz.de/passengercars/models.html>



<https://www.audi.de/de/brand/de/neuwagen.html>

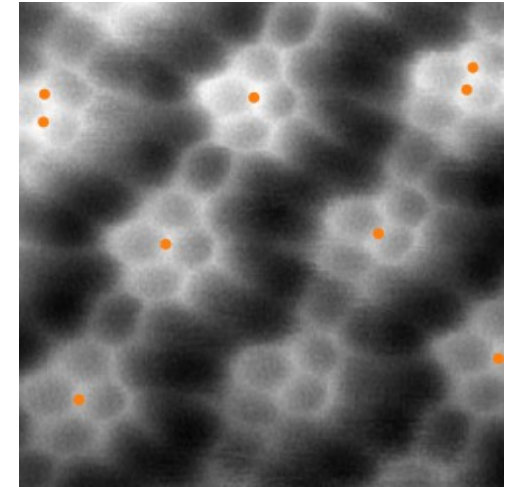


<https://www.bmw.de/de/neufahrzeuge.html>



Praxisbeispiel: Neuronale Netze in der Produktion

- Messverfahren:
kann nur Höhe erfassen, Messen der Farbe ist „physikalisch unmöglich“
- 2-tes Messsystem zum erfassen von Farbe ist angedacht
- Verlauf des Höhenprofils und Rauschen ermöglicht an einzelnen Punkten die Errechnung der Farbe via NN
- Farbinformation muss an diesen Punkte durch irgendeine Konstellation von Messartfakten vorliegen
z.B. Lichtstreuung, Erwärmung, Ausdehnung, ...
- Farbverlauf im Restlichen Bild wird durch Interpolation rekonstruiert
- Funktionsweise ist bis heute nicht klar
- Verlässlichkeit von 80% deutlich höher als notwendig



<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4050271/>

Praktisches Vorgehen

Projektbausteine

- math. Modelle
- Experimente
- Softwaremodule

Komplexes Machine Learning Model?

EVTL. OK

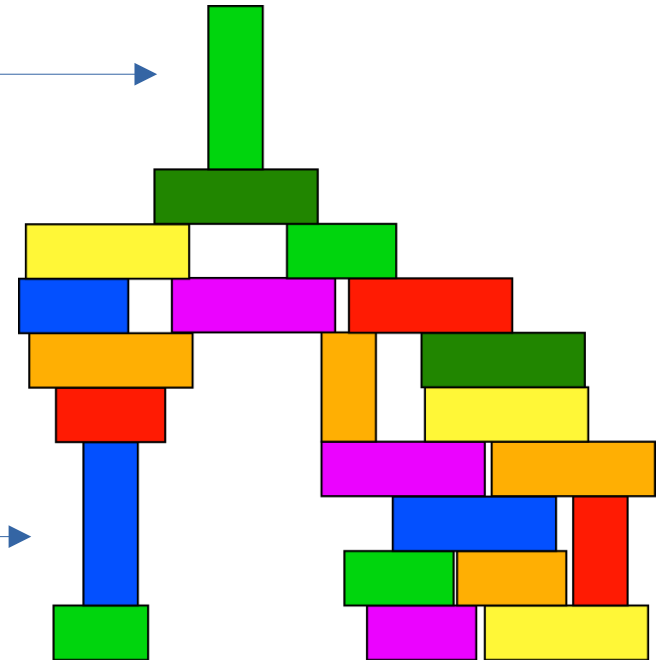
NICHT OK

Beispiel: Optische Inspektion

- Objekterkennung via NN
- Objektprüfung via Algorithmus

statt

- Objekterkennung + Prüfung via NN



Weitere Machine Learning Techniken

Natural Language Processing

- Letzte 100 Buchstaben → nächster Buchstabe

Reinforcement Learning:

- aktuelle Situation, Steuersignal → finaler Punktestand

Clustering:

- Anomalie Detektion: Weit weg von jedem Cluster

Spektakuläre Videos

- Pilco, Card-Pole: Balancieren Lernen
<https://www.youtube.com/watch?v=XiigTGKZfks>
- Boston Dynamics, BigDog: Roboterhund
<https://www.youtube.com/watch?v=cNZPRsrwumQ>
- Rotationskräfte: Einrad fahrender Roboter
<https://www.youtube.com/watch?v=RqOYLWgq2QA>
- YOLOv2: Personenerkennung umgehen
<https://www.youtube.com/watch?v=MlbFvK2S9g8>
<https://www.youtube.com/watch?v=6SXmOhTSh8Q>

Für Techies:

- DeepMind, Agent57: KI spielt auf Atari Konsole
<https://www.youtube.com/watch?v=dJ4rWhpAGFI>

Vielen Dank



Algorithmus Schmiede

Data Science | Numerik | Physik



Dr. Markus Dutschke
Inhaber

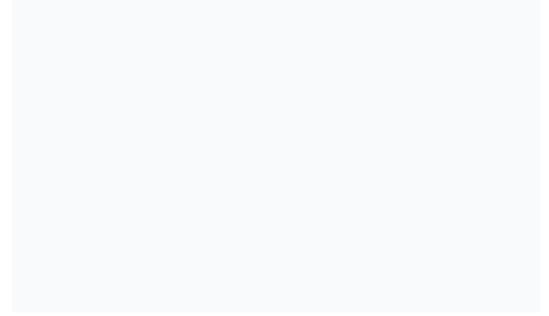
 +49 178 148 32 64

 Algorithmus Schmiede

 impact@algorithmus-schmiede.de

 www.algorithmus-schmiede.de

Backup Slides



Es folgen einige ergänzende Slides zum Vortrag.
Sollten diese Themen Ihr Interesse Wecken, wenden Sie sich gerne an uns.

Was soll schon schief gehen? Bias in Trainingsdaten

iPhone X - Face ID kann chinesische Frauen nicht voneinander unterscheiden

Aus China kommt ein Bericht, laut dem die Gesichtserkennung Face ID zwei chinesische Arbeitskolleginnen nicht voneinander unterscheiden kann.

gamestar.de

Microsoft

Twitter-Nutzer machen Chatbot zur Rassistin

Tay, ein Chatbot von Microsoft mit künstlicher Intelligenz, sollte im Netz lernen, wie junge Menschen reden. Nach wenigen Stunden musste der Versuch abgebrochen werden.

Von **Patrick Beuth**

zeit.de

24. März 2016, 15:13 Uhr / [51 Kommentare](#) / 

Video-Vorschläge von Facebook
Algorithmus verwechselt Schwarze mit Affen

04.09.2021 09:34 Uhr

zdf.de

Was soll schon schief gehen? Adversarial Examples



ALEX LEE SECURITY 11.05.2020 06:00 AM

wired.co.uk

This ugly t-shirt makes you invisible to facial recognition tech

Researchers at Northeastern University have developed an adversarial example that works even when printed onto a moving fabric

Fooling a Real Car with Adversarial Traffic Signs

Nir Morgulis, Alexander Kreines, Shachar Mendelowitz, Yuval Weisglass

Harman International, Automotive Security Business Unit arxiv.org



classification: 30 km/h

Was soll schon schief gehen?

Plausibilität, Kontext, Spezialfälle

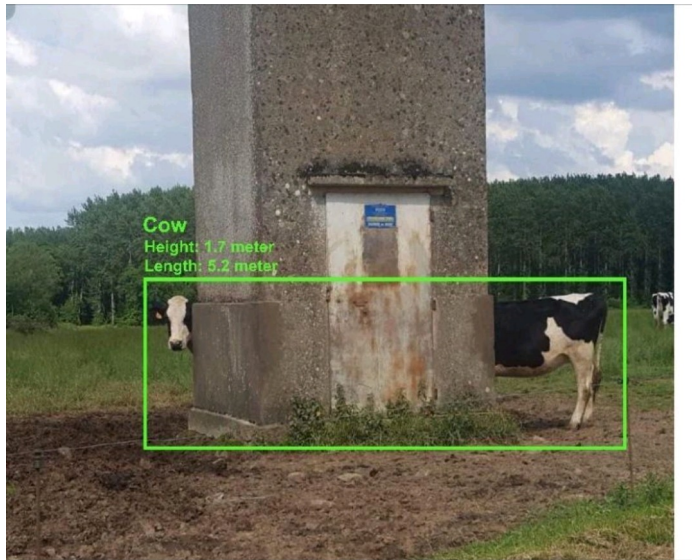


Pascal BORNET • 3rd+
LinkedIn Top Voice in Tech | ...
1mo • 🌐

+ Follow



The longest cow in the world 🤔...
Or how to deceive an AI program!... see more



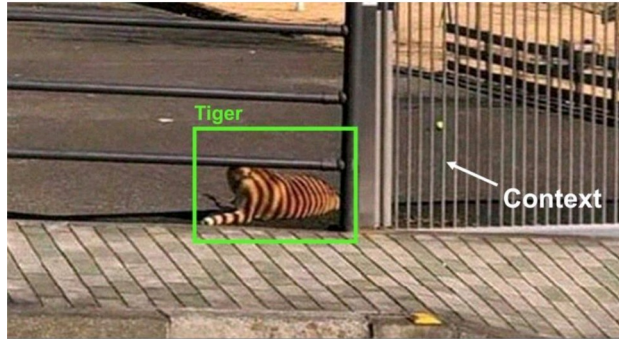
Quelle: LinkedIn



Pascal BORNET • 3rd+
LinkedIn Top Voice in Tech | Keynote ...
1d • 🌐

+ Follow

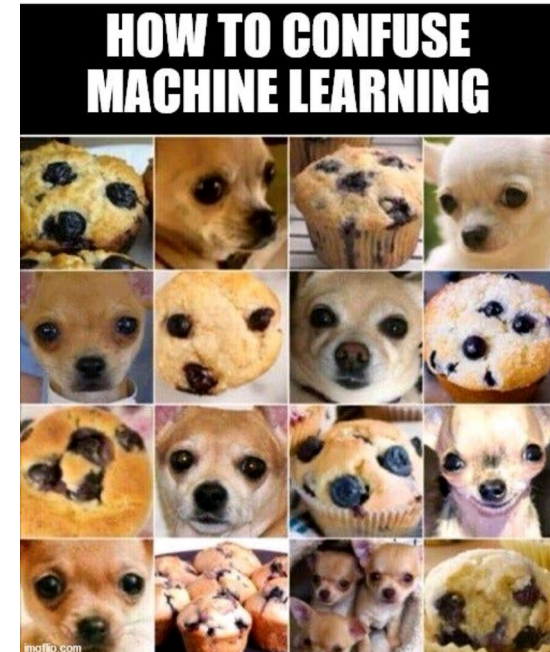
Data without context is just useless and misleading!
😄 ... see more



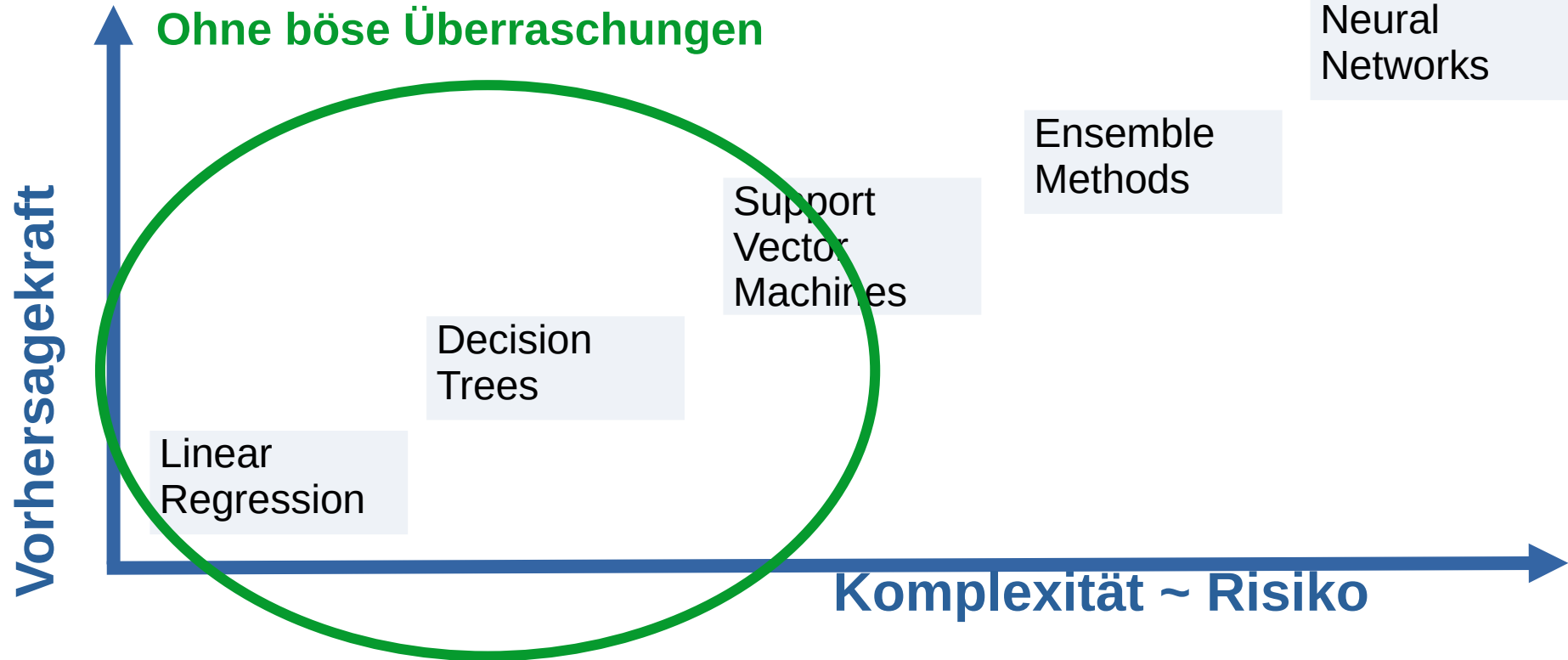
avyana
324 followers
4d • 🌐

+ Follow

How to Confuse Machine Learning... see more



Machine Learning Modelle



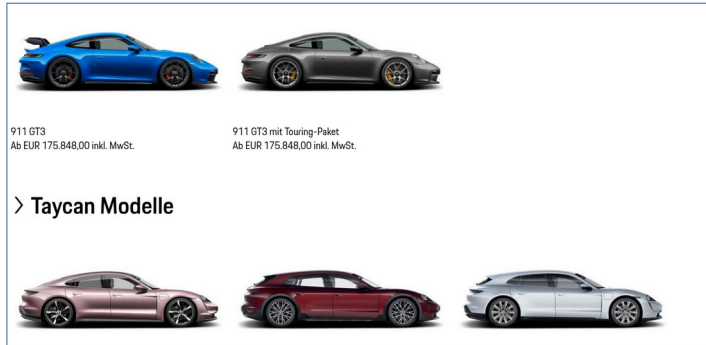
Idealfall für die Anwendungen von Neuronalen Netzen

- Daten werden automatisch erzeugt (z.B. Sensor in Produktion)
- Trainingsdaten sind repräsentativ
Temperatur: Sommer/Winter
Beleuchtung: Tag/Nach
Maschine: selbe / nicht gleiche
- Eine Lösung ist mit keiner anderen bekannten Methodik realistisch
- Ein Ausfall führt zu keinem großen Schaden oder lässt sich durch Notfallalgorithmen verhindern

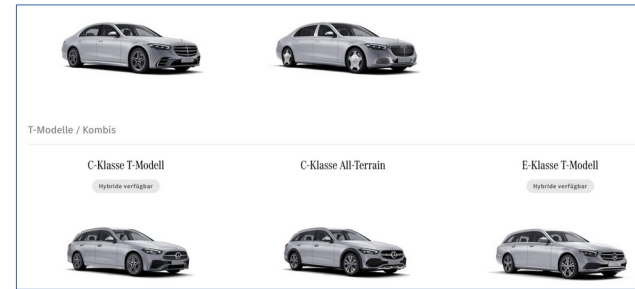
Unter diesen Umständen lässt sich der Arbeitsaufwand gut abschätzen und die Risiken halten sich in Grenzen. Je weniger der Voraussetzungen erfüllt sind, desto herausfordernder wird das Entwicklungsprojekt.

Aufgabe: Finde Automarke

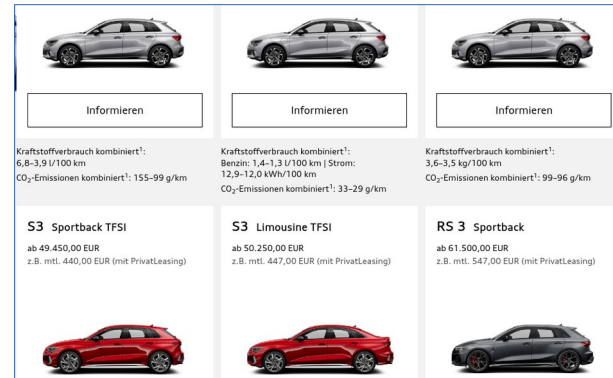
<https://www.porsche.com/germany/models/>



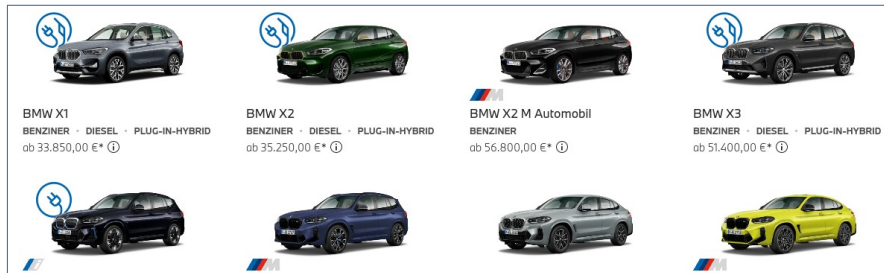
<https://www.mercedes-benz.de/passengercars/models.html>



<https://www.audi.de/de/brand/de/neuwagen.html>



<https://www.bmw.de/de/neufahrzeuge.html>



Aufgabe: Finde Automarke

Erwartete Entwicklungsstadien eines Neuronalen Netz
bis zur gewollten Funktionsweise

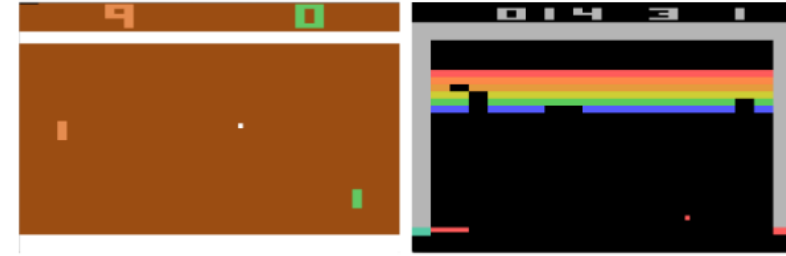
- Ausrichtung des Autos, evtl. nur untere Kante des Schattens
- Verhältnis von weißen Pixeln im Bild
- Farbe
- Metasymbole  
- Winkel der Beleuchtung / Schattenverläufe entlang der Karosserie
- ...
- Nebensächlichkeiten wie: Anhängerkupplung, Blinker im Seitenspiegel, Felgen, ...
- ...
- Modellreihen zu jeder Marke + relevante Merkmale wie Herstelleremblem

Reinforcement Learning: Anwendungen

- Komplexe Aufgaben in der Regelungstechnik
- Alle Probleme ähnlich zu Gesellschafts- oder Computergames
- Atari 2600 Spiele

Beispiele:

- Objekte greifen und stapeln
- Spurhalte-Assistent
- Balance Aufgaben
- Notfallsteuerung, um Produkt (z.B. Lebensmittel) nach Systemfehler noch zu Retten
- Routenplanung in dynamisch veränderlicher Umgebung
- Dynamische Preisgestaltung



Reinforcement Learning

Funktionsweise:

- Regressor: (aktuelle Situation, Steuersignal) -> finaler Punktestand
- Training des Regressors etwas trickreich

Sobald der Regressor fertig ist, ist das Problem gelöst, denn:

- Berechne den finalen Punktestand für jedes mögliche Steuersignal in der aktuellen Situation.
- Wähle das Steuersignal, welches in der aktuellen Situation den höchsten finalen Punktestand erzielt.

Reinforcement Learning

Anwendung: Selbstlernender Yahtzee / Kniffel Bot

- Erreicht nach 8000 Spielen übermenschliche Spielstärke

