
Überblick: Anwendung von Maschine Learning in der Glücksspielindustrie

Dr. Jochen Fiedler

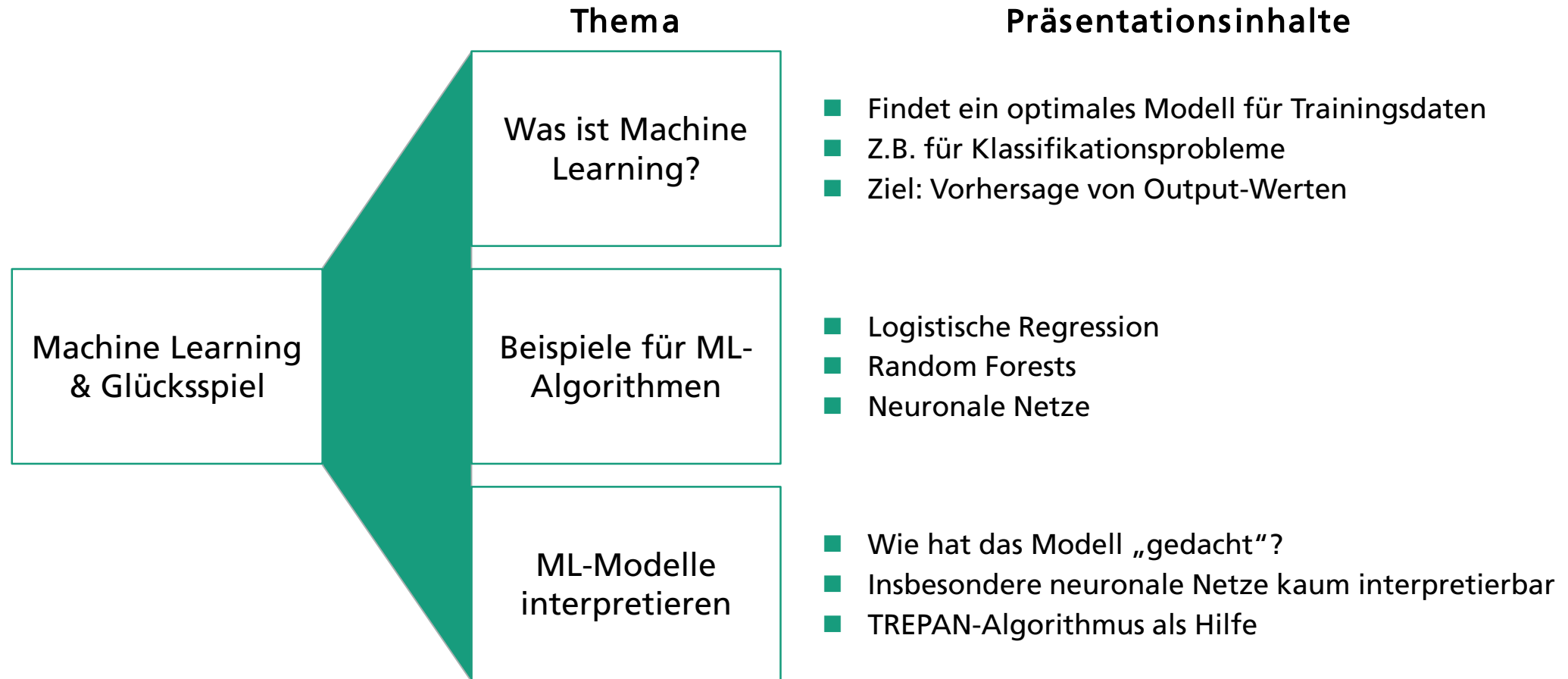
Dr. Sascha Feth

19. Symposium Glücksspiel

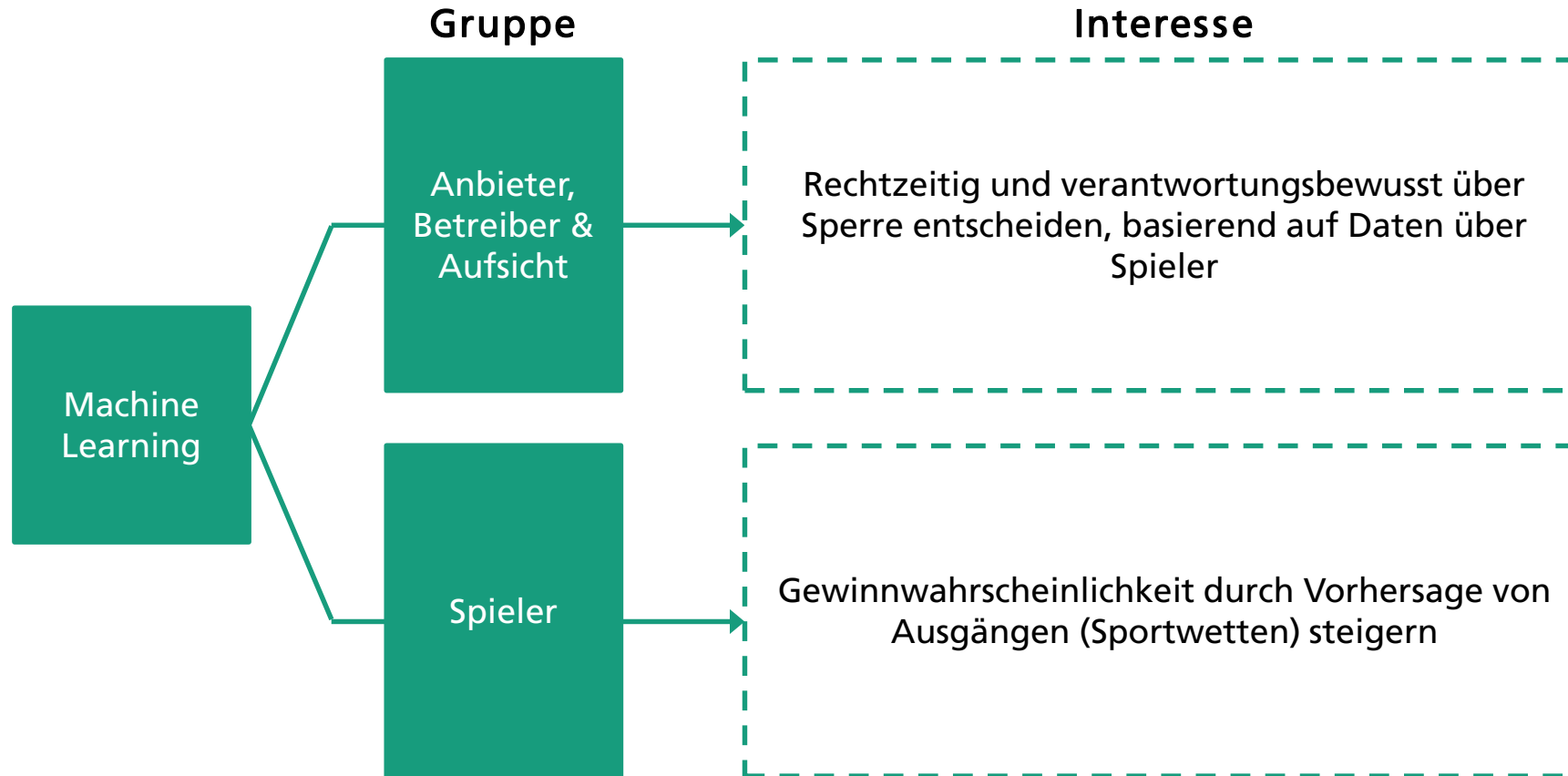
15.-16. März 2022



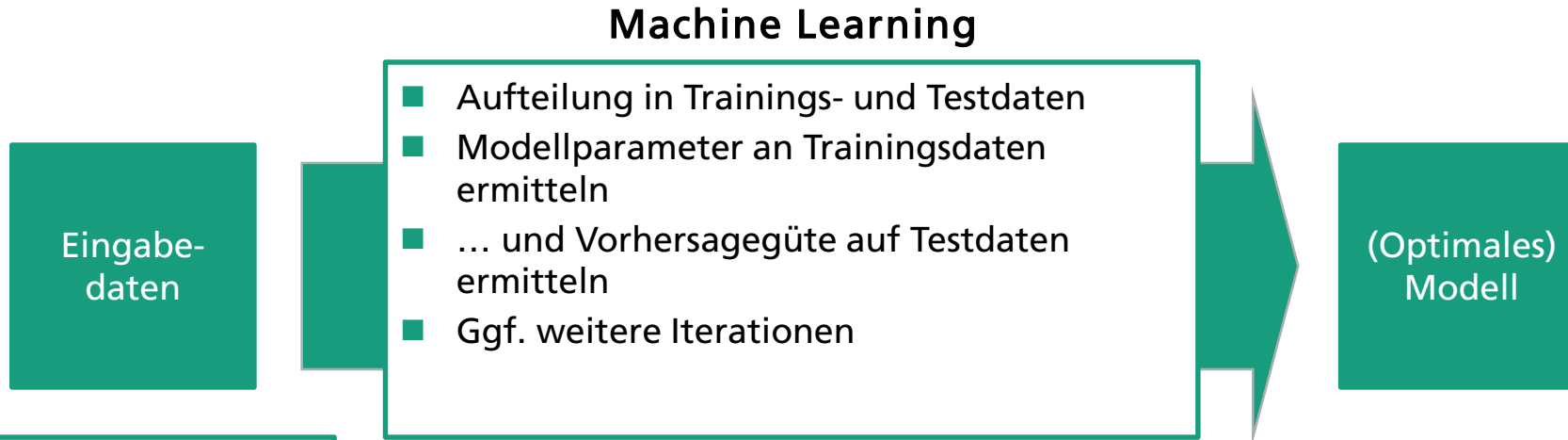
Inhalt



Machine Learning ist für Anbieter, Betreiber und Spieler interessant



Maschine Learning (ML) erstellt ein optimales Modell aus Eingabedaten



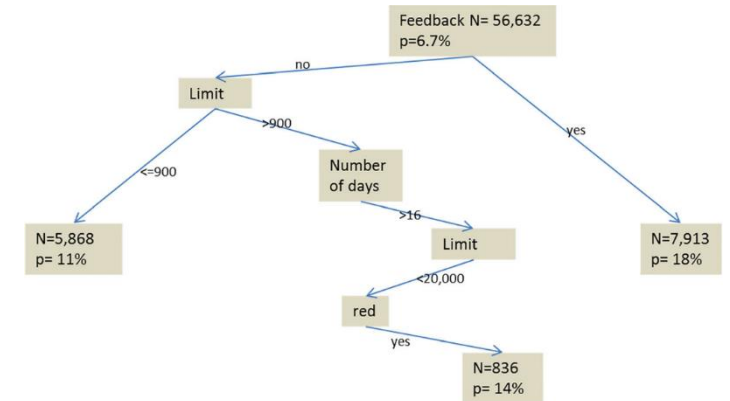
Datenpool:

- 33 Variablen von knapp 71000 norwegischen Spielern
- 2 Kategorien: Selbstausschluss, kein Selbstausschluss

Variablen:

- Einsatz, Feedback, Limit, Verlust, Möglicher Verlust, Gewinn, Persönliche Daten,...

Güte (Training)	Güte (Test)
99%	73%

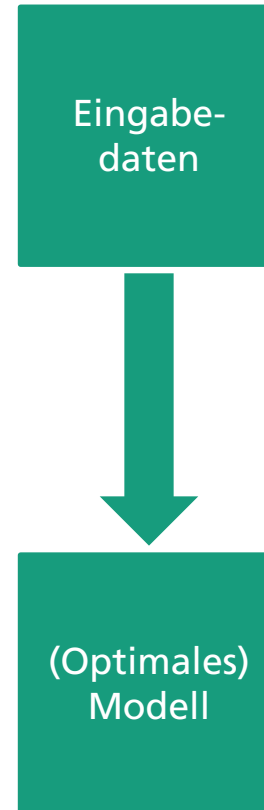


Das optimale Modell ist nicht immer offensichtlich

Beispiel „Pizza“

Teig	Belag	Käse	Schmeckt?
Dick	Salami	Normal	Ja
Dick	Pilze	Extra	Nein
Dünn	Salami	Normal	Ja
Dünn	Pilze	Extra	Nein
Dünn	Salami	Extra	Ja

Salami => Pizza schmeckt
(ziemlich) offensichtlich

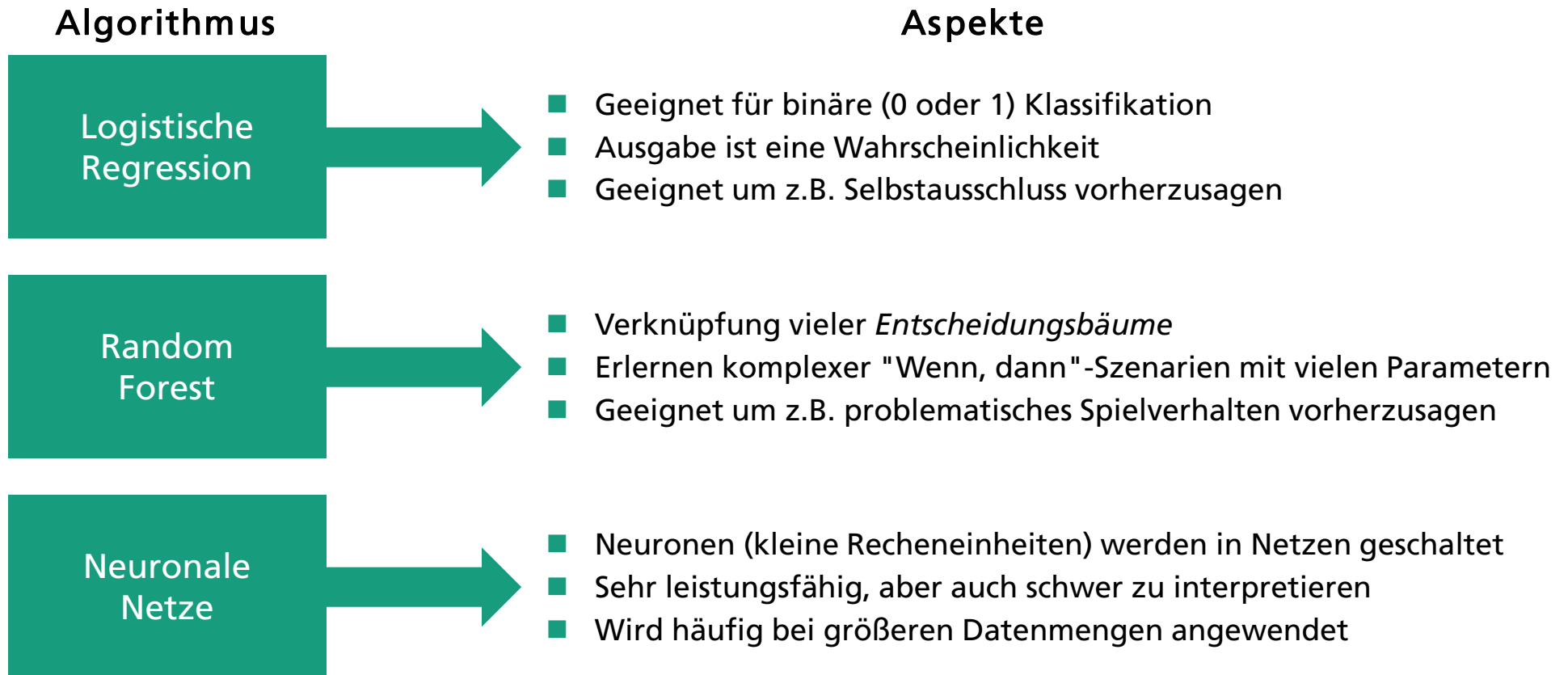


Beispiel „Pasta“

Form	Teig	Soße	Schmeckt?
Farfalle	Normal	Tomate	Ja
Fusilli	Normal	Tomate	Nein
Farfalle	Vollkorn	Tomate	Ja
Fusilli	Vollkorn	Käse	Ja
Farfalle	Normal	Käse	Ja

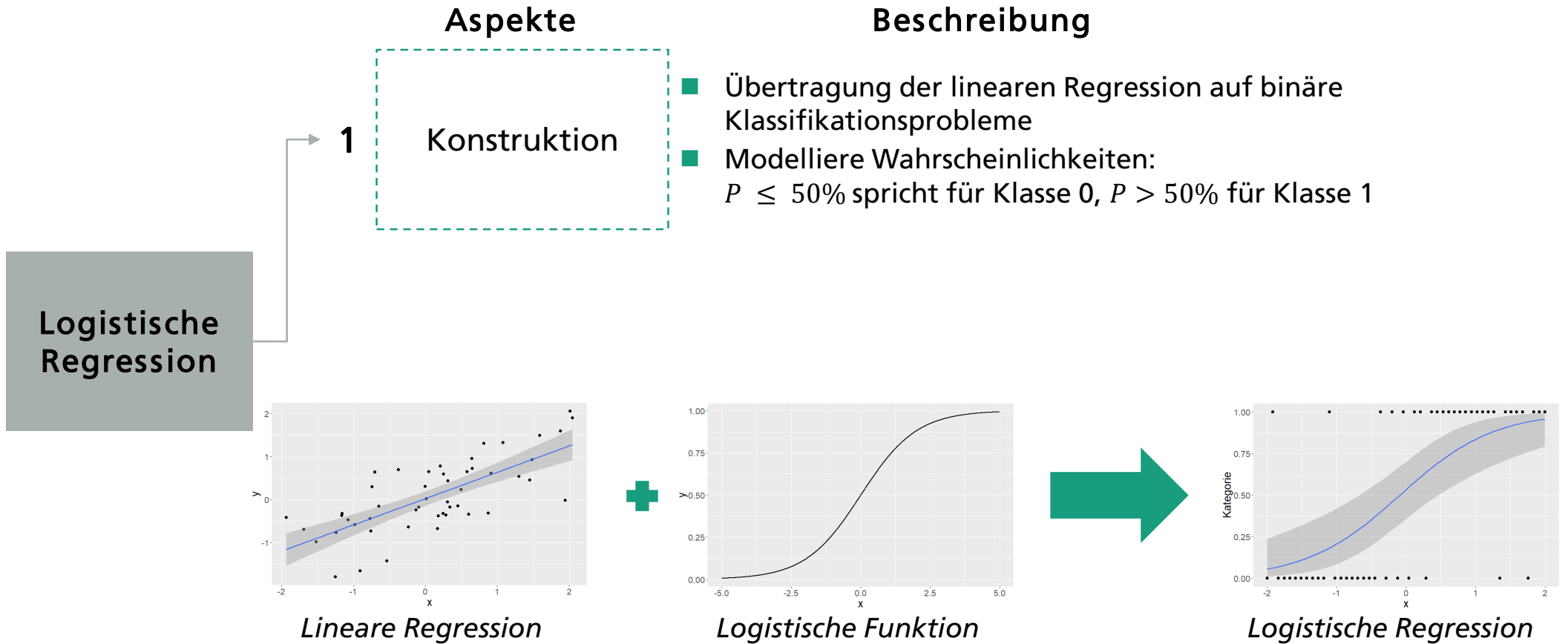
??? => Pasta schmeckt
(nicht) offensichtlich

Wir stellen exemplarisch drei ML-Algorithmen vor (Anwendungsbeispiele auf letzter Folie zitiert)

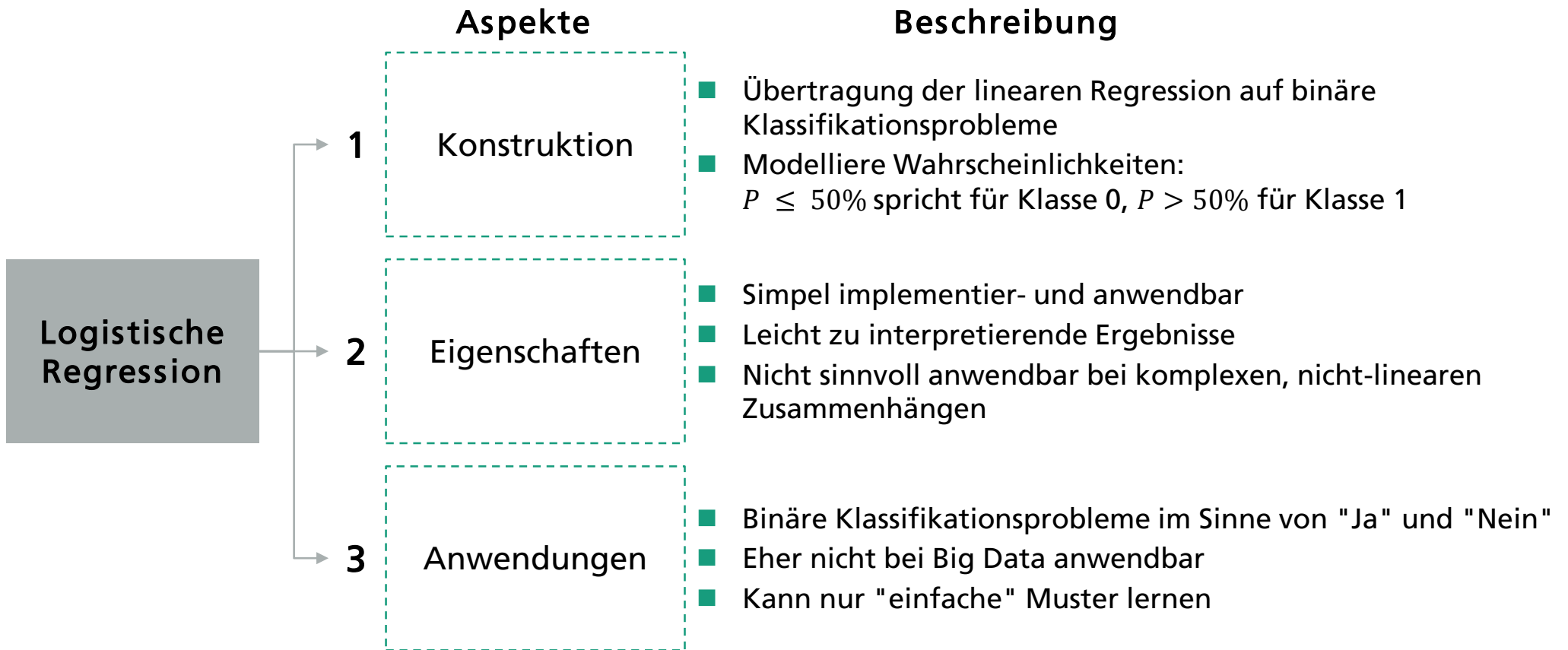


LOGISTISCHE REGRESSION

Logistische Regression überträgt klassische Regression auf Klassifikationsprobleme



Logistische Regression überträgt klassische Regression auf Klassifikationsprobleme



Ein konstruiertes Beispiel

Datenpool:

- 2 Variablen von 3000 fiktiven Spielern
- 2 Kategorien: Selbstausschluss, kein Selbstausschluss

Variablen:

- Monatliches Budget
- Monatliche Spielzeit

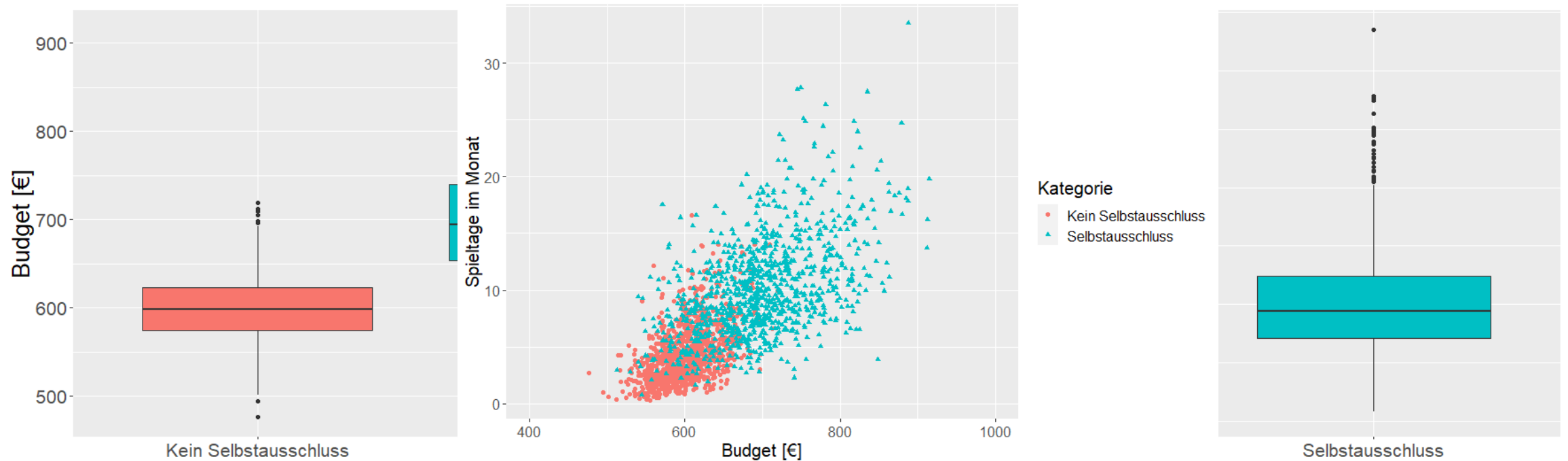
Ziel:

- Vorhersage ob Selbstausschluss oder nicht



Ein konstruiertes Beispiel

Spieler in den verschiedenen Kategorien unterscheiden sich systematisch



Ein konstruiertes Beispiel

Datenpool:

- 2 Variablen von 3000 fiktiven Spielern
- 2 Kategorien: Selbstausschluss, kein Selbstausschluss

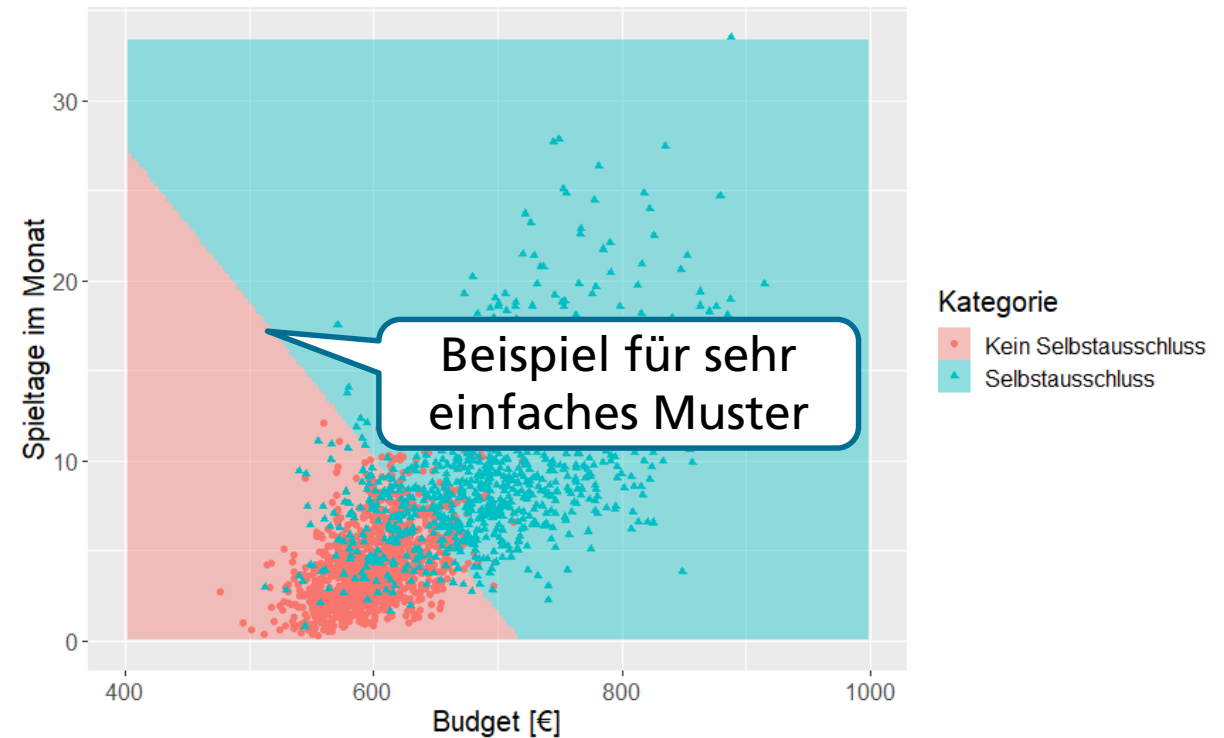
Variablen:

- Monatliches Budget
- Monatliche Spielzeit

Ziel:

- Vorhersage ob Selbstausschluss oder nicht

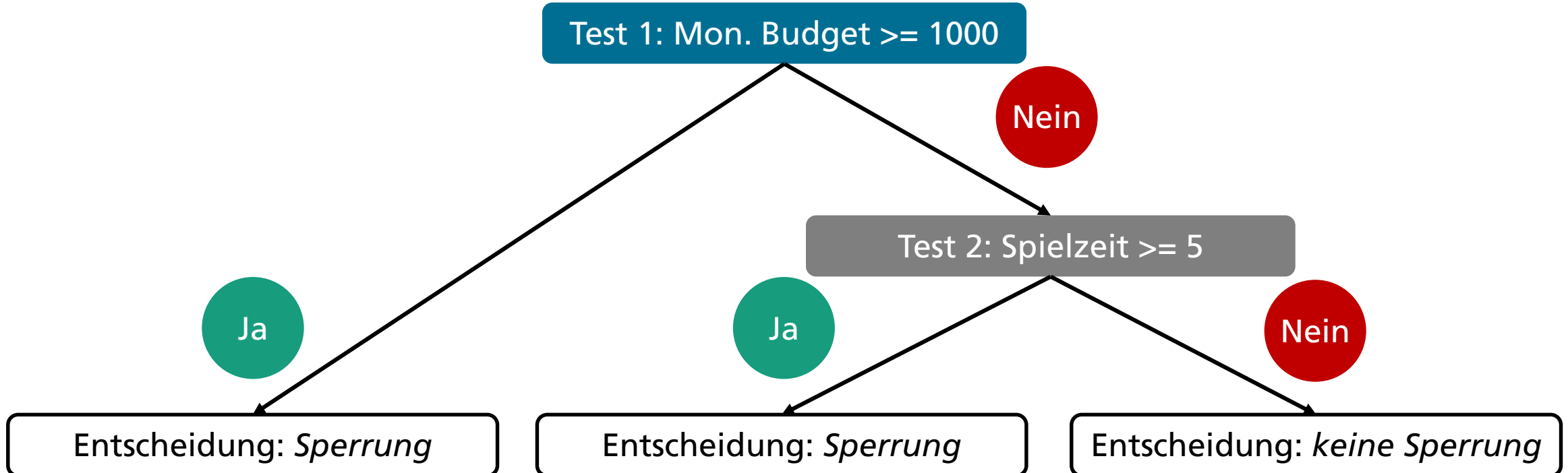
Güte (Training)	Güte (Test)
86%	84%



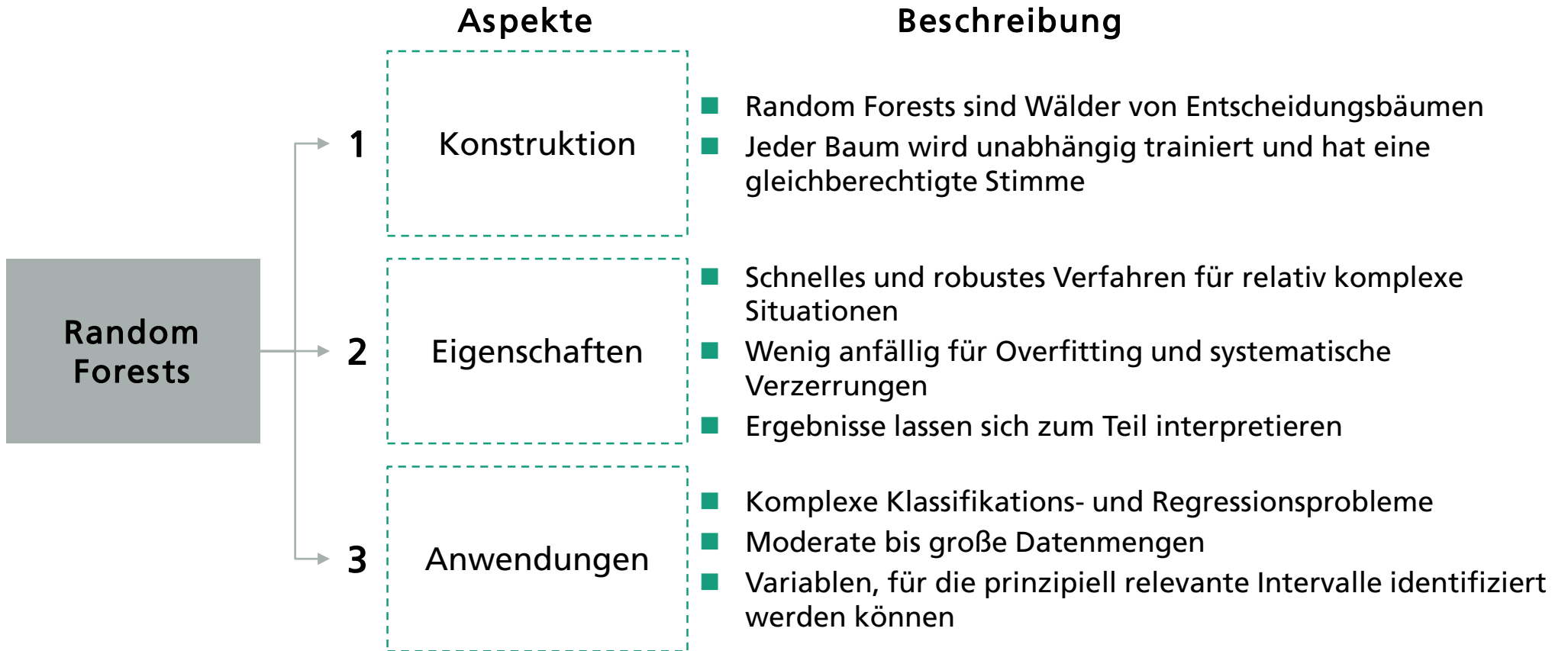
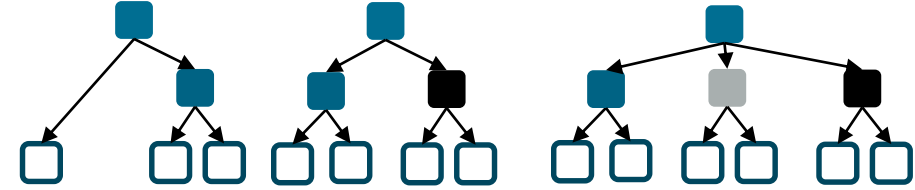
RANDOM FOREST

Klassifikationsbäume bilden die Grundlage von Random Forests

Idee: Führe sukzessive Tests anhand einzelner Erklärungsvariablen durch, bis es zu einer Entscheidung kommt



Random Forests bestehen aus vielen Entscheidungsbäumen



Anwendungsbeispiel: Setzen von Limits von Spielern

Studie von Auer und Griffiths (2019) auf Basis von Daten norwegischer Spieler

Datenpool:

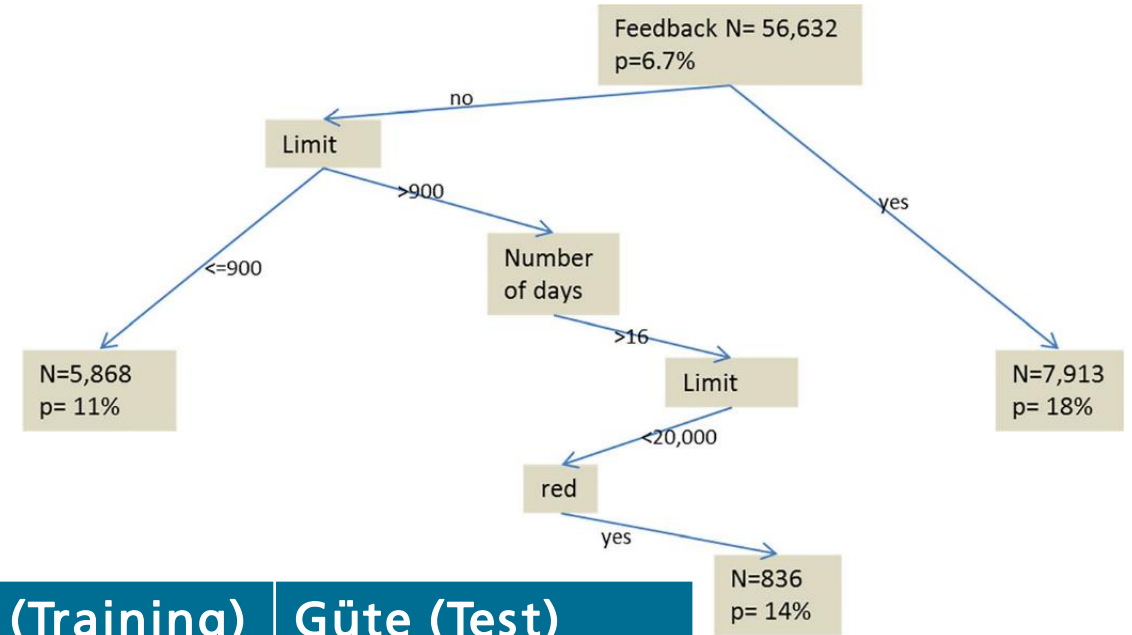
- 33 Variablen von knapp 71000 norwegischen Spielern
- 2 Kategorien: Limit geändert, Limit nicht geändert

Variablen:

- Einsatz, Feedback, Limit, Verlust, Möglicher Verlust, Gewinn, Persönliche Daten,...

Ziel:

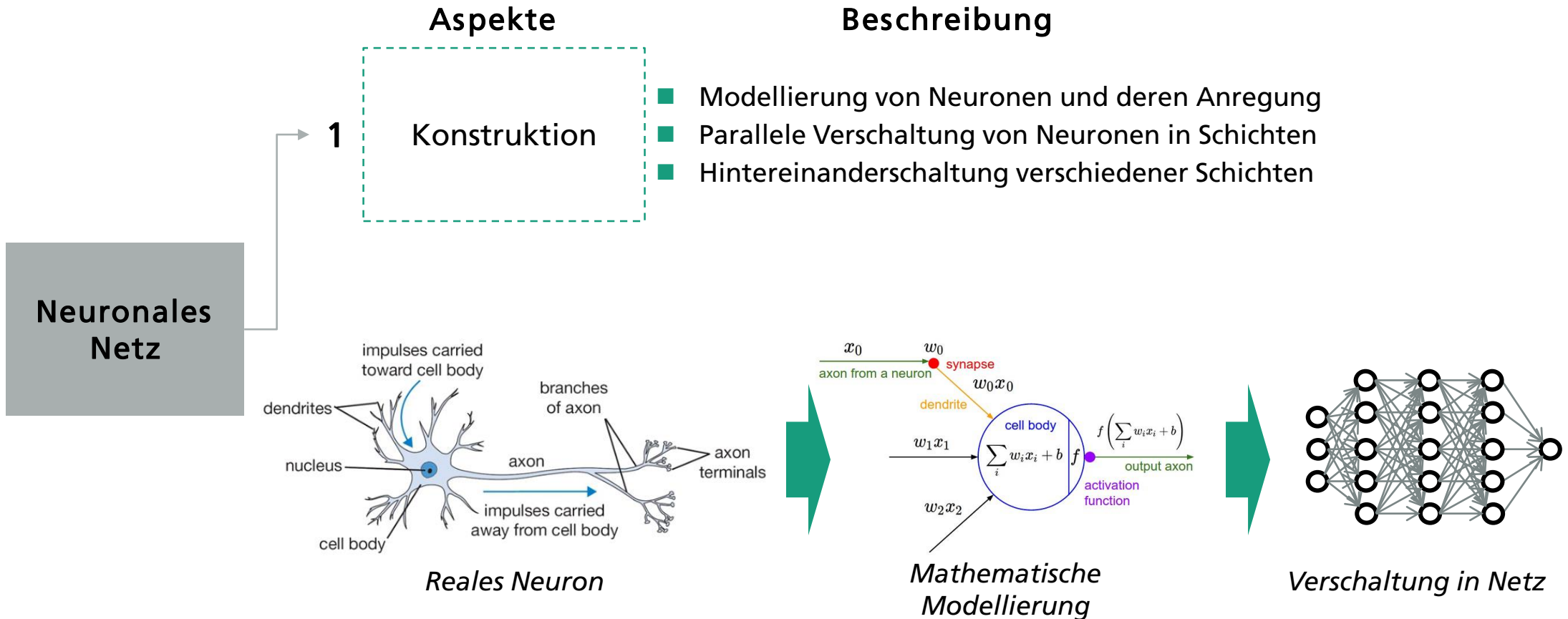
- Vorhersage ob Limit geändert wurde



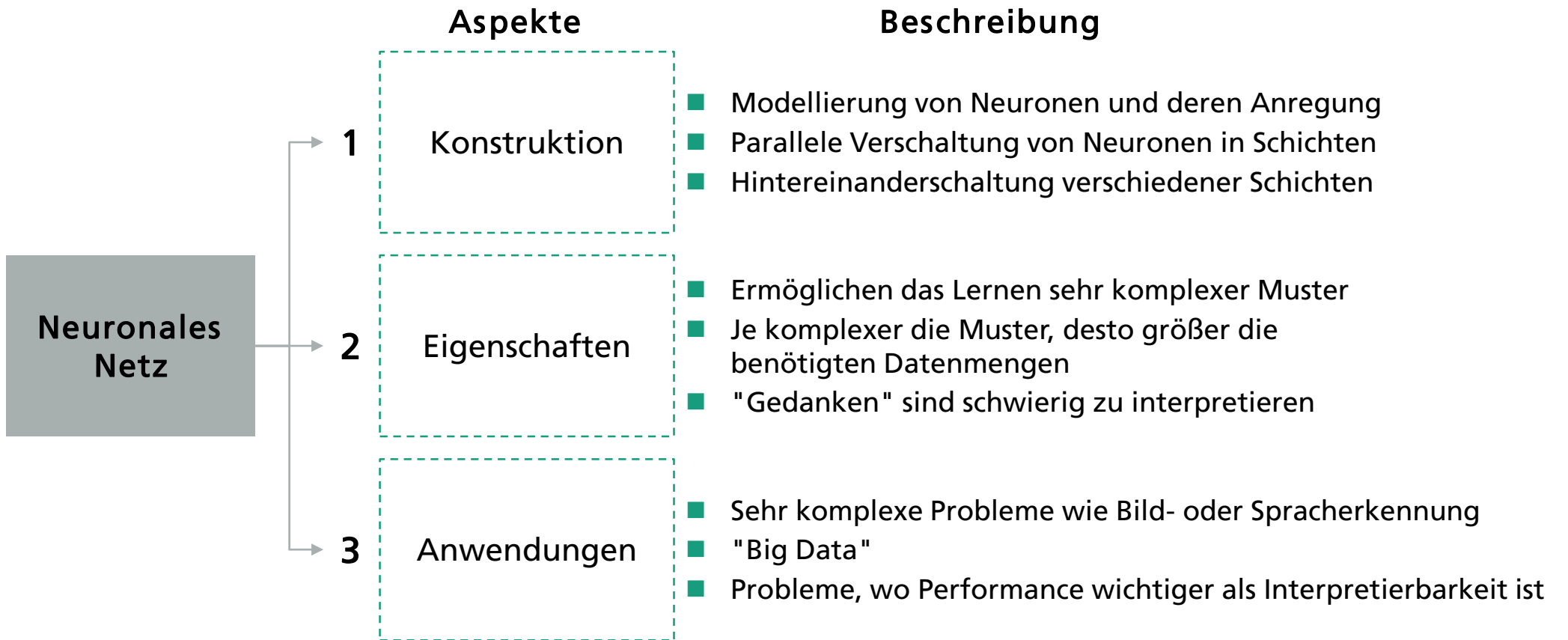
Güte (Training)	Güte (Test)
99%	73%

NEURONALE NETZE

Neuronale Netze orientieren sich an Gehirnfunktionen



Neuronale Netze orientieren sich an Gehirnfunktionen



Anwendungsbeispiel: Schädliches Spielverhalten

Percy et al. (2016) untersuchen schädliches Spielverhalten anhand von Selbstausschluss

Datenpool:

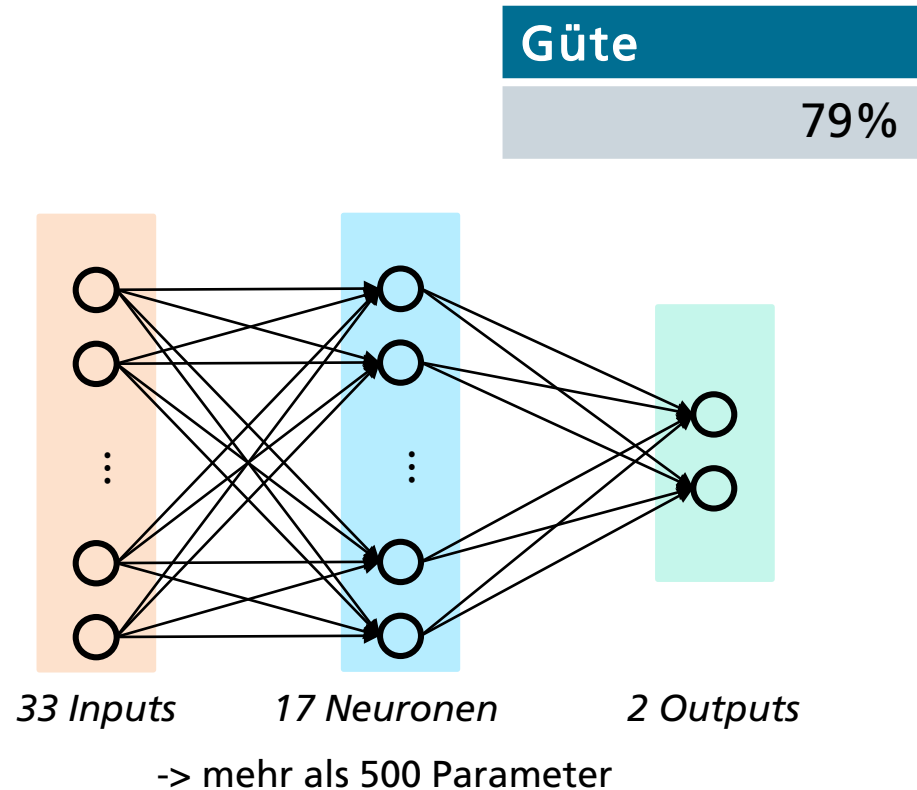
- *International Game Technology (IGT)*
Daten von 845 Onlinespielern
- 2 Kategorien: *Selbstausschluss* (min. 6 Monate) und *kein Selbstausschluss*

Variablen:

- Abgeleitet aus Rohdaten, bspw. Gesamteinsatz über den Tag, Variabilität beim Setzen,...

Ziel:

- Vorhersage ob Selbstausschluss oder nicht



INTERPRETIERBARKEIT



Quelle: Wikipedia (Trepanation)

ML-Modelle sind für Menschen schwierig zu interpretieren

ML-Modelle

Random Forests

Neuronale Netze

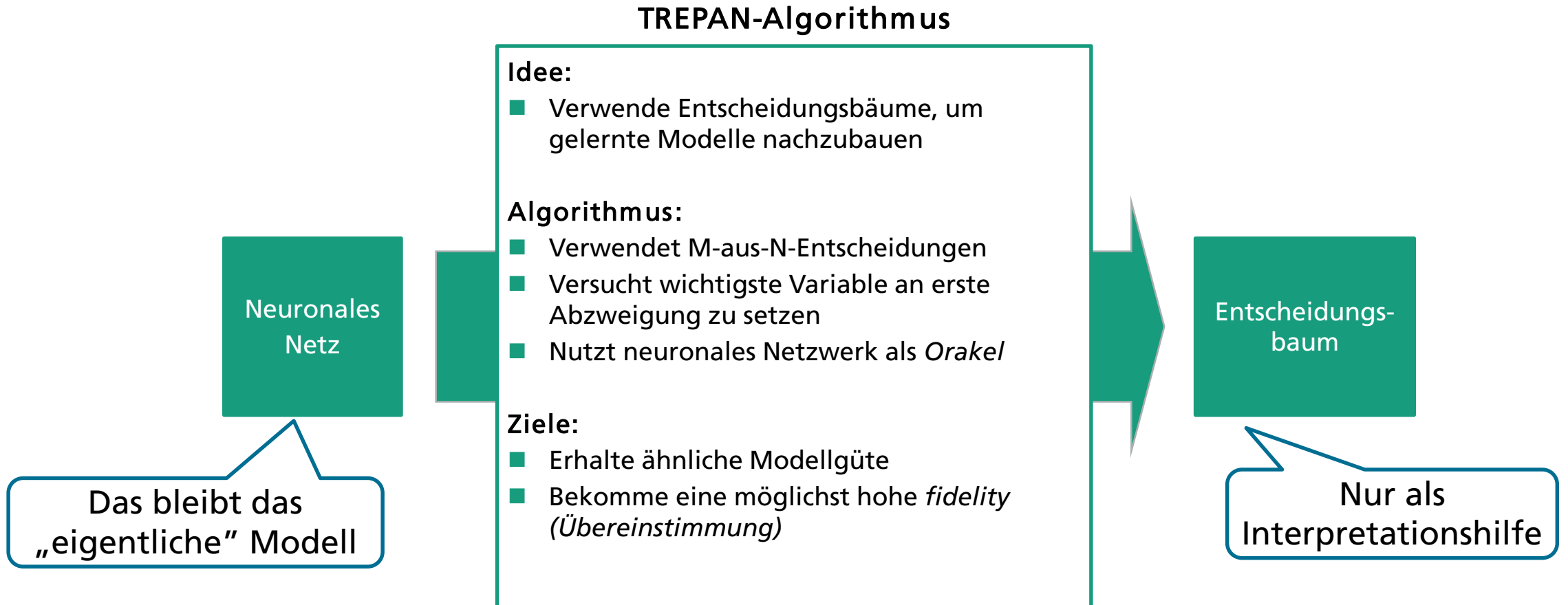


Interpretation

- Die Wichtigkeit einzelner Variablen lässt sich ableiten
 - Einzelne Entscheidungsbäume können nachvollzogen werden – bei der Gesamtheit ist es schwieriger
- Lässt sich nicht in einfache "Wenn, dann"-Entscheidungen übersetzen

- Neuronale Netze bestehen aus einigen Schichten mit vielen Neuronen
 - Dies ergibt eine große Menge von gelernten Gewichten, welche alle Informationen enthalten
- Nicht von Menschen direkt interpretierbar

Der TREPAN-Algorithmus erleichtert die Interpretation

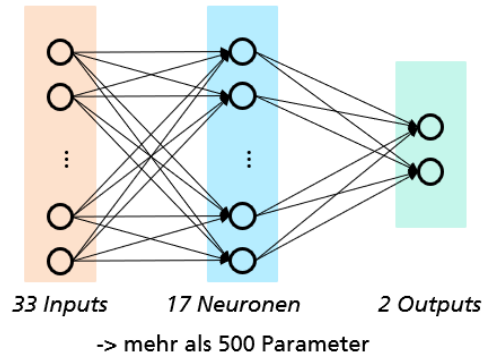


Anwendungsbeispiel für den TREPAN-Algorithmus

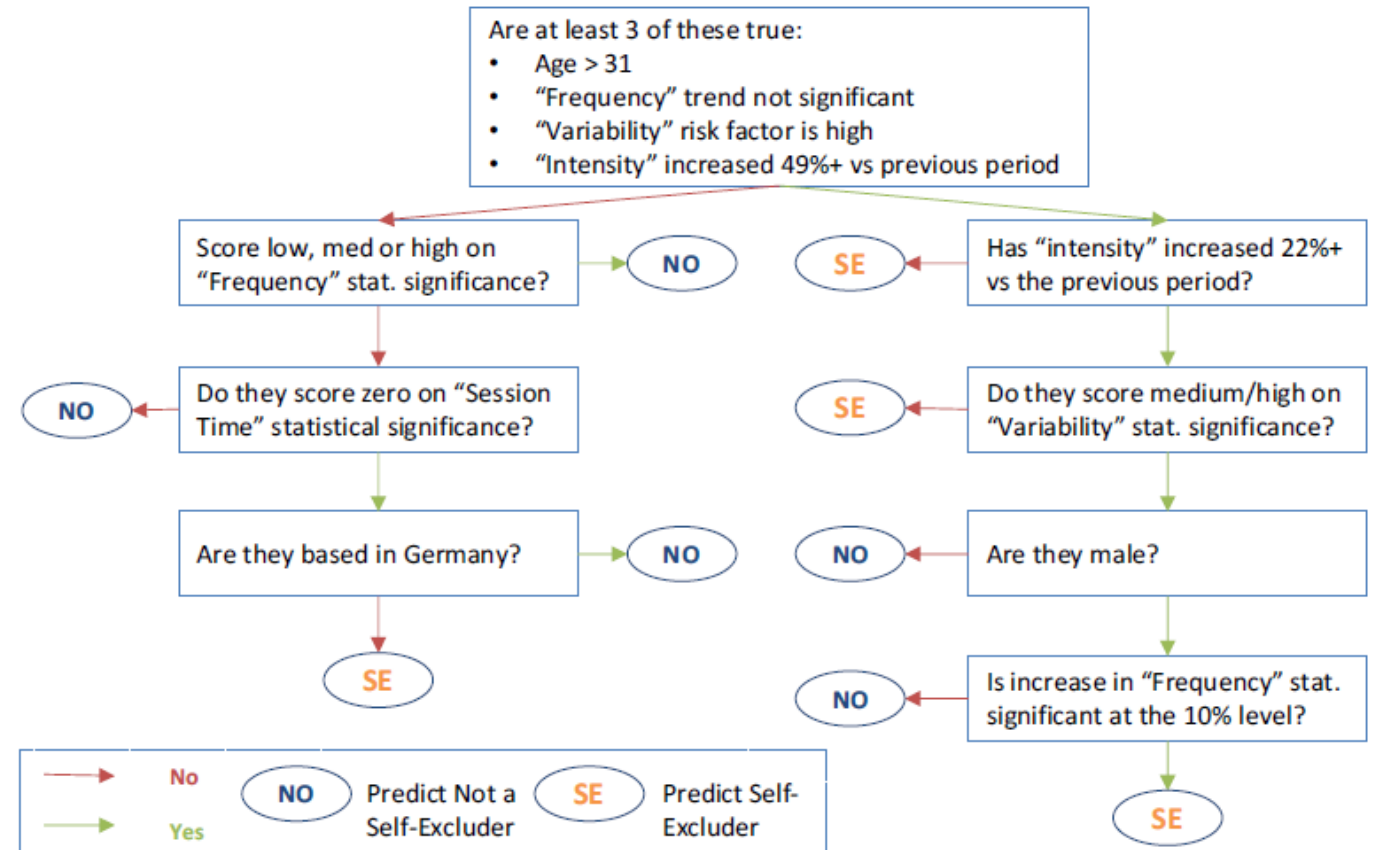
Situation:

- Von Percy et al. (2016) angewendet
- Ziel: Schädliches Spielverhalten erklären

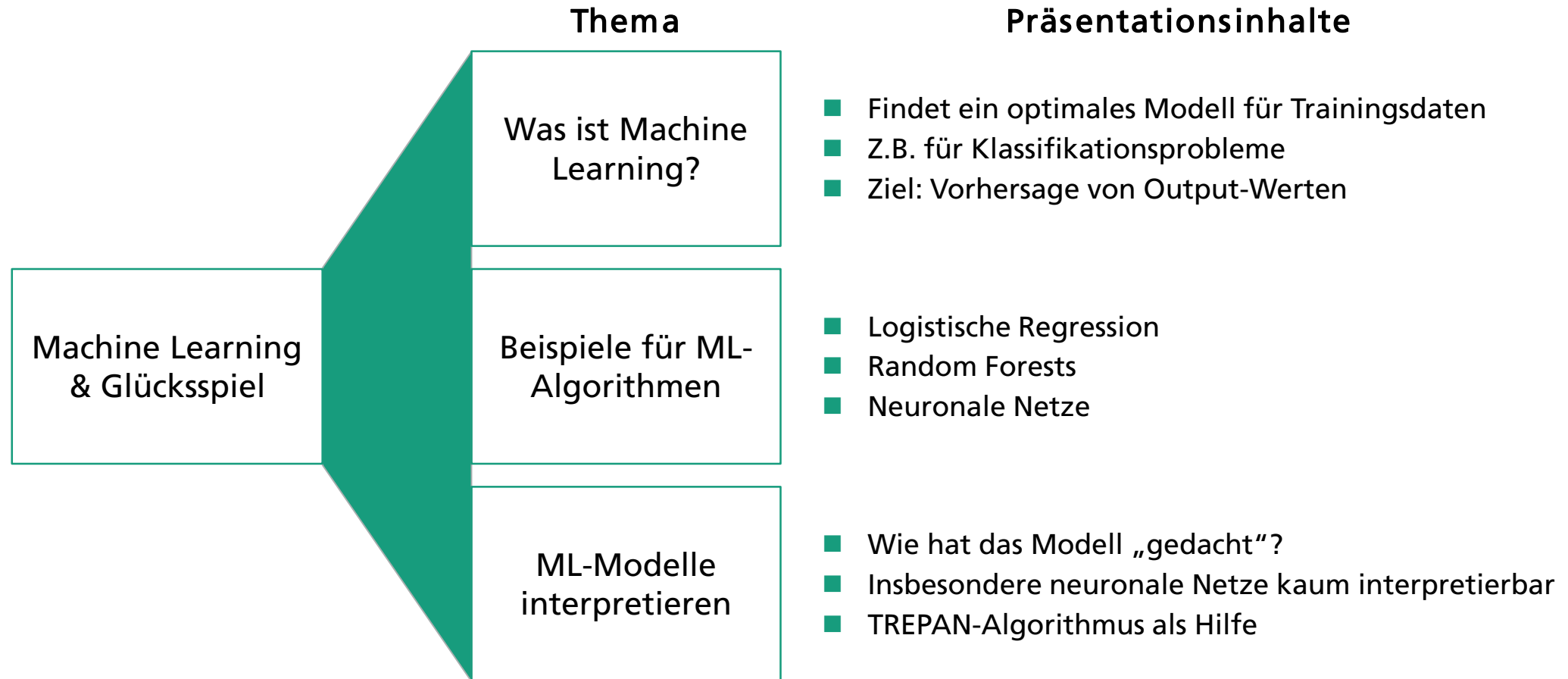
Resultate:



Modell	Güte	Fidelity
Neuronales Netz	79,8%	-
TREPAN	78,8%	87,4%



Zusammenfassung



Quellen

■ Artikel in einer Zeitschrift:

- [1] Auer, Michael und Griffiths, Mark D.: *Predicting Limit-Setting Behavior of Gamblers Using Machine Learning Algorithms: A Real-World Study of Norwegian Gamblers Using Account Data*, in: **International Journal of Mental Health and Addiction** (2019)
- [2] Craven, M and Shavlik, J.: *Extracting Tree-Structured Representations of Trained Networks*, in: **Advances in Neural Information Processing Systems** (1996), 37-45
- [3] Percy, C.; d'Avila Garcez, A. S.; Dragičević, S.; França, M. V. M.; Slabaugh, G. and Weyde, T.: *The Need for Knowledge Extraction: Understanding Harmful Gambling Behavior with Neural Networks*, in: **ECAI** (2016)

■ Podcast & Blog „Streuspanne“

- <https://www.itwm.fraunhofer.de/de/abteilungen/mf/aktuelles/blog-streuspanne.html>
- Folge 3: „Gewinnen Statistiker:innen häufiger im Glücksspiel?“