

# ABSCHLUSSVORTRAG

SENSORBASIERTER ORIENTIERUNGSSINN  
MIT KÜNSTLICHEN NEURONALEN NETZEN  
UND ENTSCHEIDUNGSBÄUMEN

---

Tom Dymel

16.06.2021

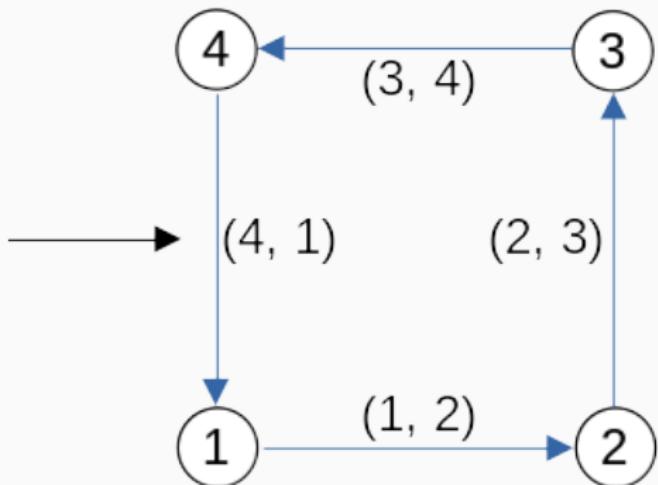
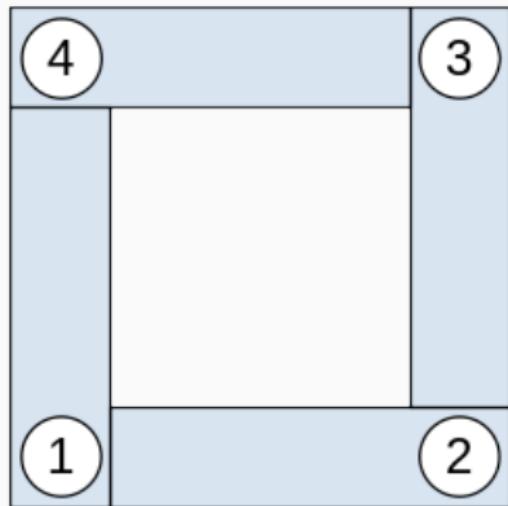
Masterarbeit

Technische Universität Hamburg

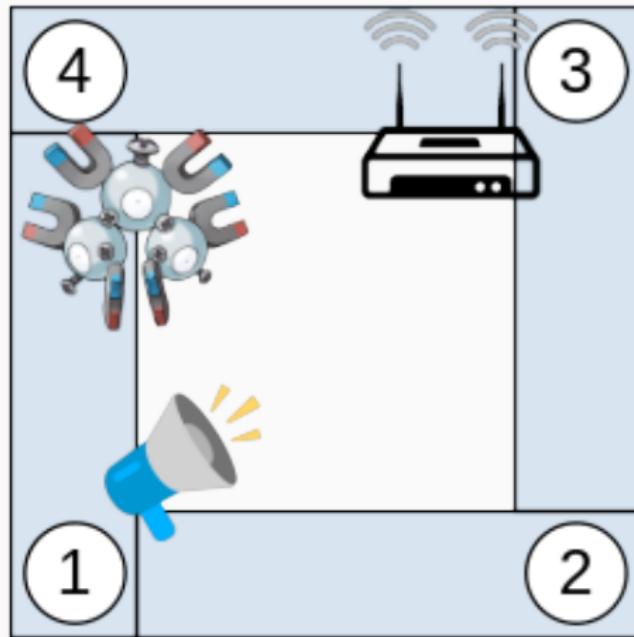
# MOTIVATION UND ZIELE

- Motivation
  - Indoor-Lokalisierung hohe Infrastrukturkosten
  - Orientierungssinn von Tieren und Menschen
  - Mian untersuchte FFNN und simulierte Daten
  - Entscheidungsbäume potentiell effizienter
- Ziele
  - Diskrete Standortbestimmung
  - Nutzung von mehreren Sensoren
  - Ausgewählte Features bewerten

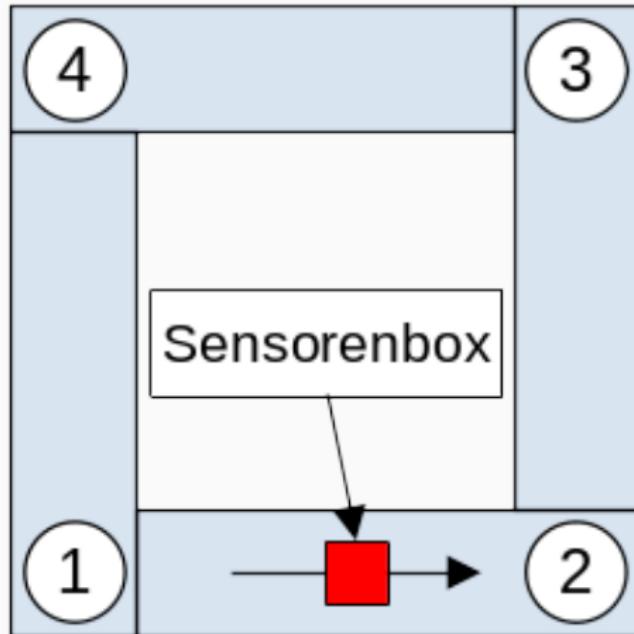
# ANSATZ ZUR DISKREten STANDORTBESTIMMUNG I



# ANSATZ ZUR DISKREten STANDORTBESTIMMUNG II

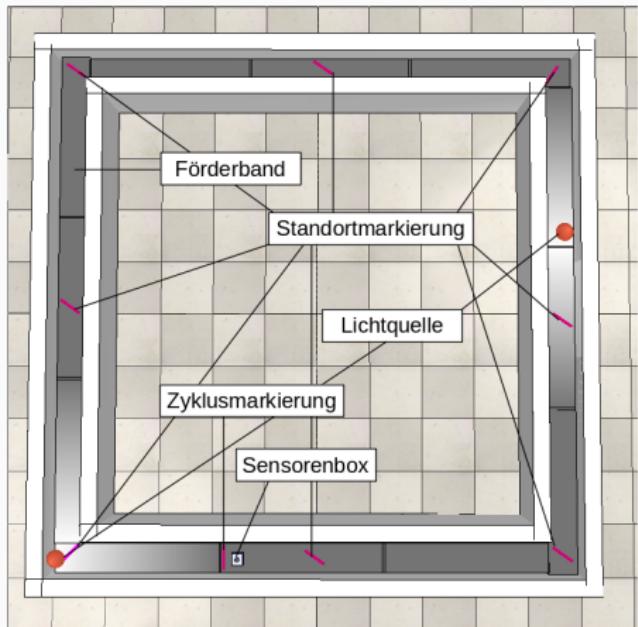


# ANSATZ ZUR DISKREten STANDORTBESTIMMUNG III

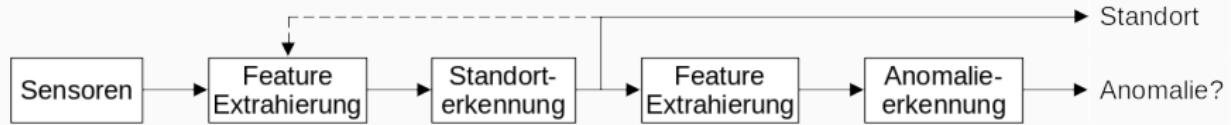


# SENSOREN

- Accelerometer
- Gyroskop
- Licht
- Magnetfeld
- Temperatur
- Geräusche
- WLAN-Zugangspunkte



# MODELL

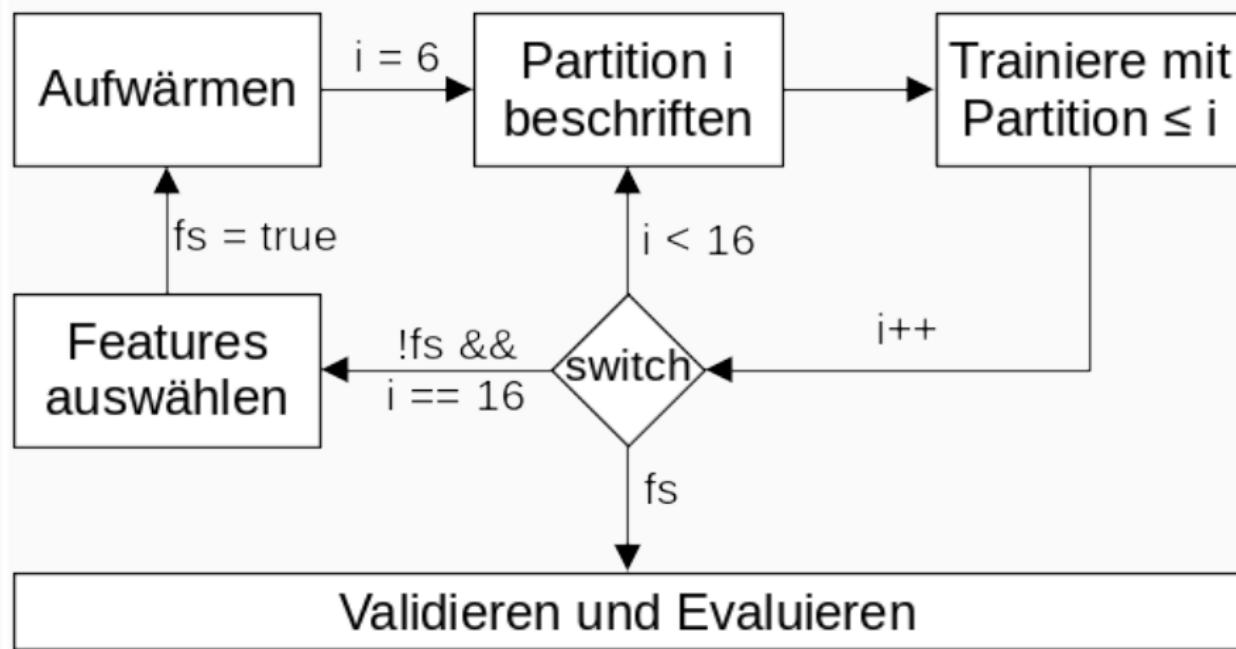


# FEATURE-EXTRAHIERUNG

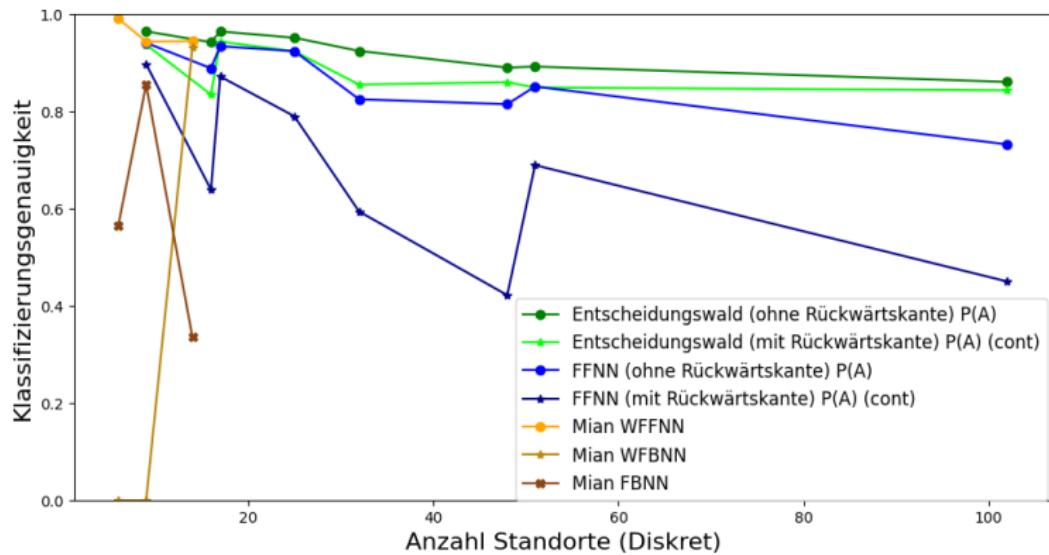
| Sensordaten                        | $\sigma$ | Min. | Max. | $\emptyset$ | Wert |
|------------------------------------|----------|------|------|-------------|------|
| Accelerometer                      | X        | X    | X    | X           | X    |
| Gyroskop                           | X        | X    | X    | X           | X    |
| Ausrichtung zum Magnetfeld         | X        | X    | X    | X           | X    |
| Temperatur                         | X        | X    | X    | X           | X    |
| Licht                              | X        | X    | X    | X           | X    |
| Geräusch                           | X        | X    | X    | X           | X    |
| WLAN-Zugangspunkte                 | -        | -    | -    | -           | X    |
| Letzter Standort                   | -        | -    | -    | -           | X    |
| Letzter unterschiedlicher Standort | -        | -    | -    | -           | X    |
| Zeit                               | X        | -    | -    | -           | -    |

# TRAINING DER ML-MODELLE

|                   |           |   |            |
|-------------------|-----------|---|------------|
| Partition i:      | 1 - 5     | 6 - 15  | 16 - 20    |
| Phase:            | Aufwärmen | Beschriften mit Modell                        | Validieren |
| Beschriftet in %: | 0         | $100 \cdot \min(1, 0.5 + (i - 6)^2 \div 128)$ | -          |



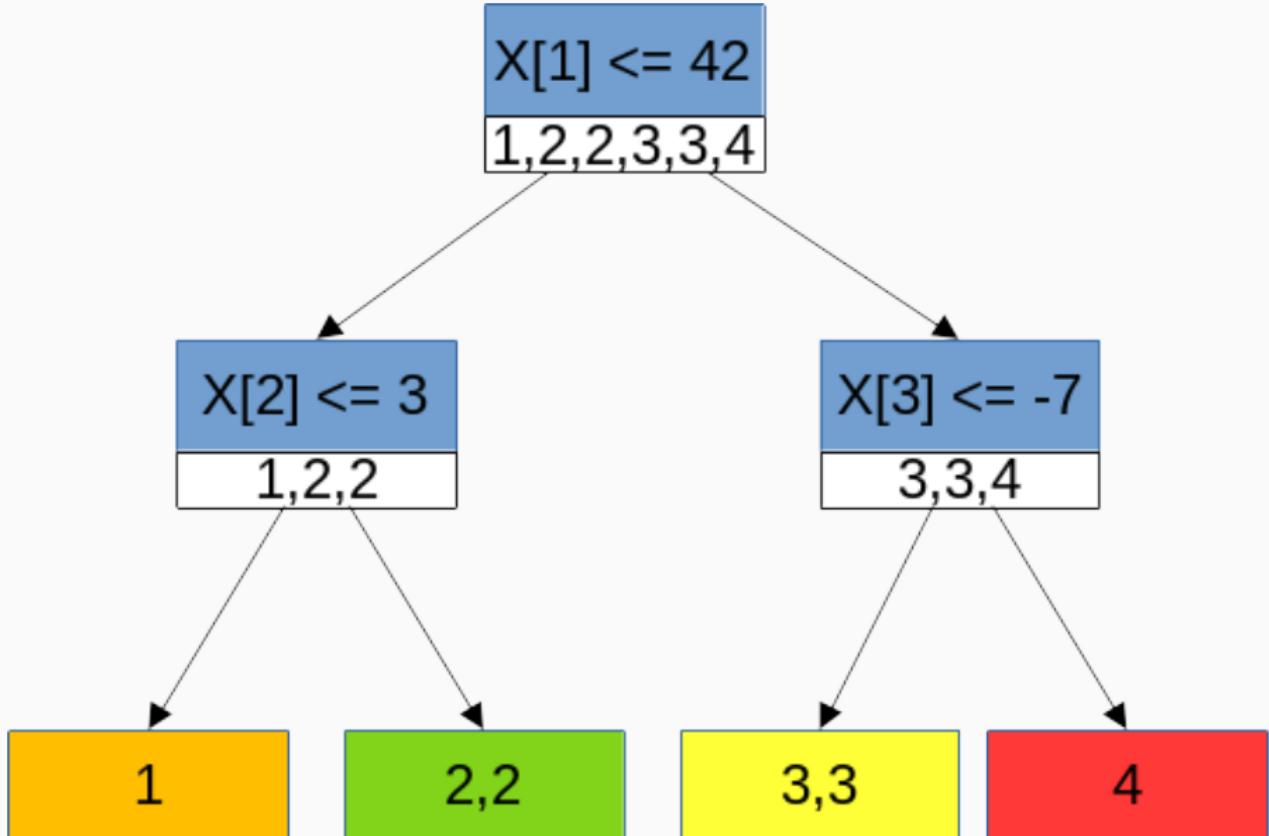
# KLASSIFIZIERUNGSGENAUIGKEIT ÜBER STANDORTKOMPLEXITÄTEN



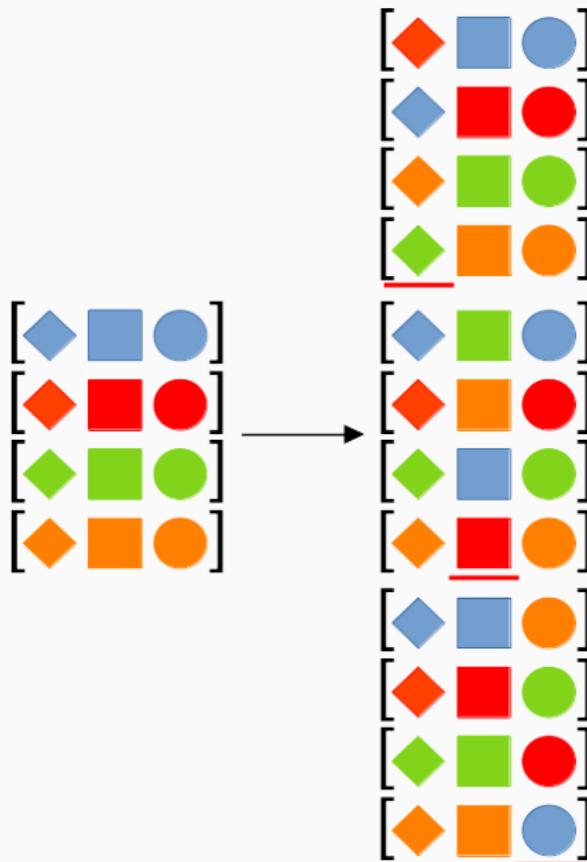
# WICHTIGKEIT VON FEATURES I

1. Basierend auf Entscheidungsregeln und Trainingsdaten
2. Modifizieren der Testmenge und Fehler zur Originalmenge messen, z. B. Permutationswichtigkeit oder Nullung

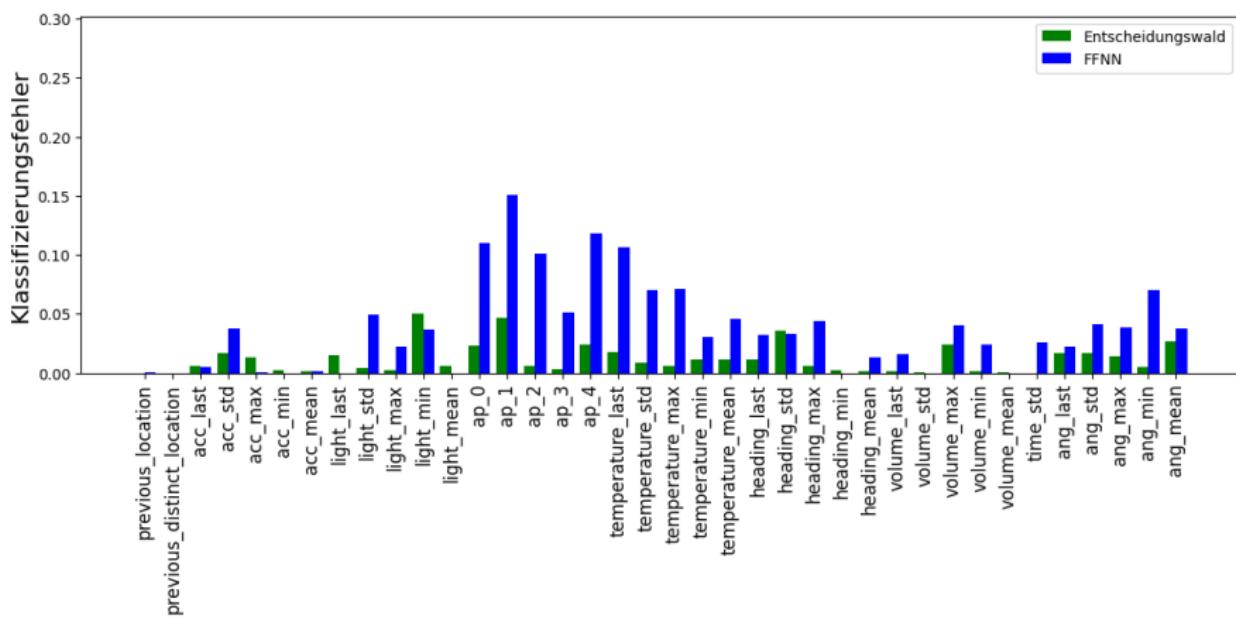
## WICHTIGKEIT VON FEATURES II - feature\_importances\_



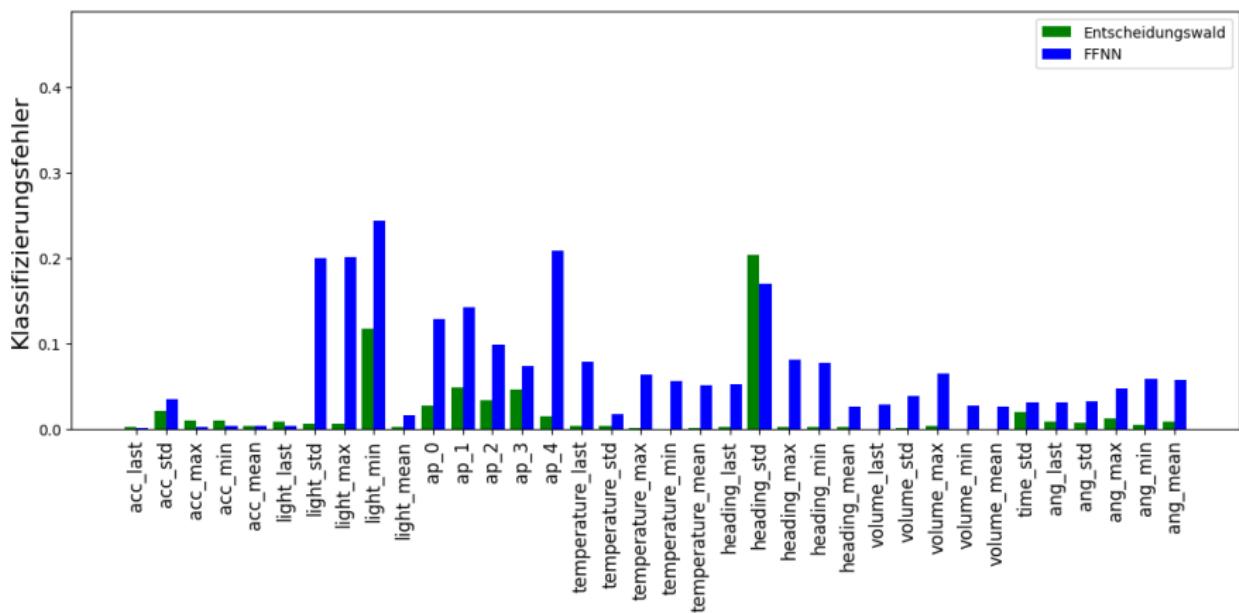
## WICHTIGKEIT VON FEATURES III - PERMUTATIONSWICHTIGKEIT



# PERMUTATIONSWICHTIGKEIT - ML-MODELLE MIT RÜCKWÄRTSKANTE



# PERMUTATIONSWICHTIGKEIT - ML-MODELLE OHNE RÜCKWÄRTSKANTE



# FEHLERTOLERANZ

| Testmenge                  | Entscheidungswald  | FFNN        | Entscheidungswald   | FNNN        |
|----------------------------|--------------------|-------------|---------------------|-------------|
|                            | mit Rückwärtskante |             | ohne Rückwärtskante |             |
| Licht                      | 4.46%-Pkt.         | 4.65%-Pkt.  | 5.28%-Pkt.          | 6.93%-Pkt.  |
| Geräusch                   | 3.20%-Pkt.         | 5.00%-Pkt.  | 1.63%-Pkt.          | 5.11%-Pkt.  |
| Temperatur                 | 15.15%-Pkt.        | 6.60%-Pkt.  | 8.10%-Pkt.          | 13.50%-Pkt. |
| Ausrichtung zum Magnetfeld | 3.32%-Pkt.         | 19.94%-Pkt. | 2.51%-Pkt.          | 2.78%-Pkt.  |
| WLAN-Zugangspunkte         | 2.60%-Pkt.         | 22.65%-Pkt. | 3.74%-Pkt.          | 14.13%-Pkt. |
| Accelerometer              | 1.41%-Pkt.         | 9.52%-Pkt.  | 0.62%-Pkt.          | 1.33%-Pkt.  |
| Gyroskop                   | 8.52%-Pkt.         | 4.58%-Pkt.  | 0.91%-Pkt.          | 3.30%-Pkt.  |
| Permutierte Testmenge      | 2.27%-Pkt.         | -0.13%-Pkt. | 0.47%-Pkt.          | 0.93%-Pkt.  |
| Durchschnitt               | 5,8%-Pkt.          | 9,1%-Pkt.   | 2,91%-Pkt.          | 6,00%-Pkt.  |

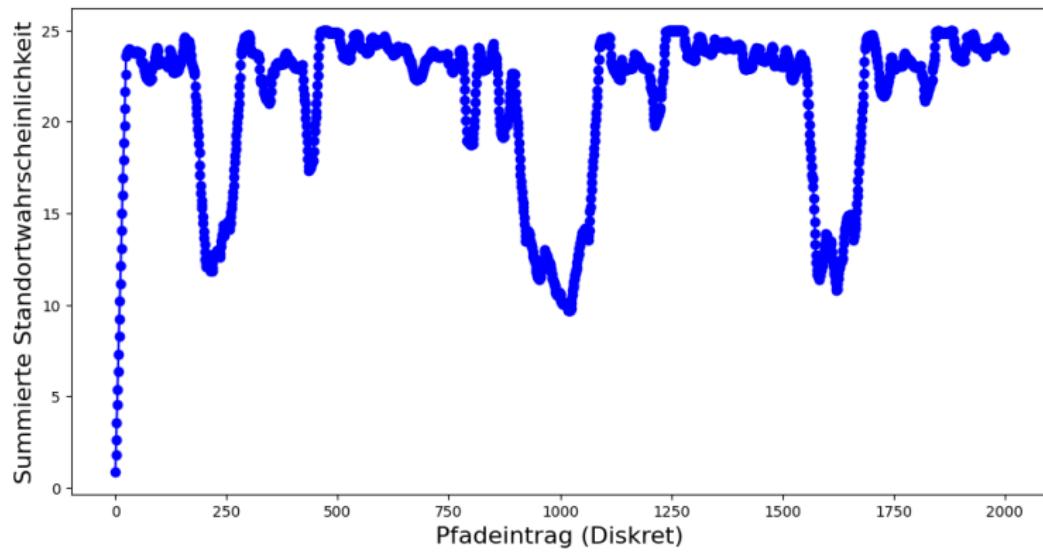
## RESSOURCENNUTZUNG

- Entscheidungswälder benötigen deutlich mehr Programmspeicher
- Entscheidungswälder benötigen deutlich weniger RAM
- Entscheidungswälder sind deutlich effizienter
- Programmgröße skaliert mit Standortkomplexität
- FFNNs zwischen 54% und 97,6% kleiner als Mians FFNNs
- Entscheidungswälder bis zu 720% größer als Mians FFNNs
- Für beide ML-Modelle existieren Mikrocontroller mit genug RAM und Programmspeicher

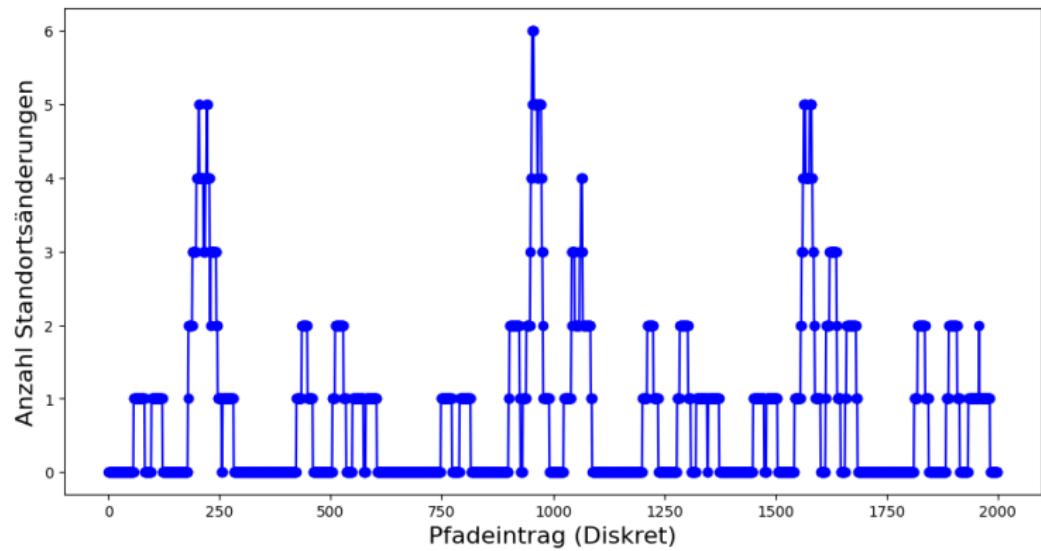
# ANOMALIEERKENNUNG

- Klassifizierung, dass momentaner Standort nicht trainiert wurde
- Unendliche viele mögliche Anomalien
- Training basierend auf Sensordaten nicht sinnvoll
- Klassifizierung ungewöhnlichen Verhaltes des Standorterkennungsmodells

# SUMMIERTE KLASIFIZIERUNGSGENAUIGKEIT IN DATENFENSTER



# ANZAHL STANDORTÄNDERUNGEN IN DATENFENSTER



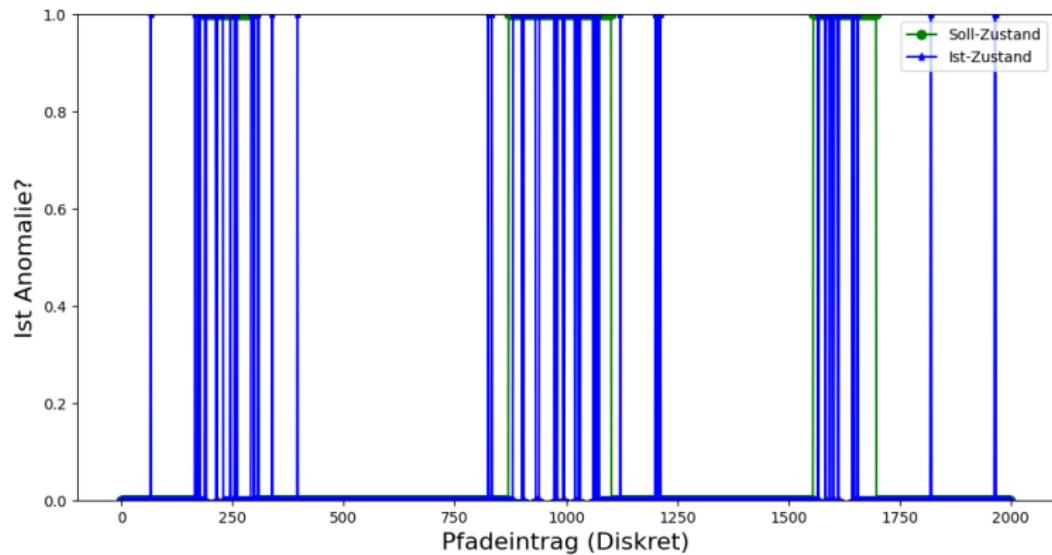
# FEATURE-EXTRAHIERUNG

- Abweichung zu  $\emptyset$  Standortänderungen
- Abweichung zu  $\emptyset$  Standortwahrscheinlichkeit
- Standardabweichung Top 5 Standortwahrscheinlichkeiten
- Topologieverletzung

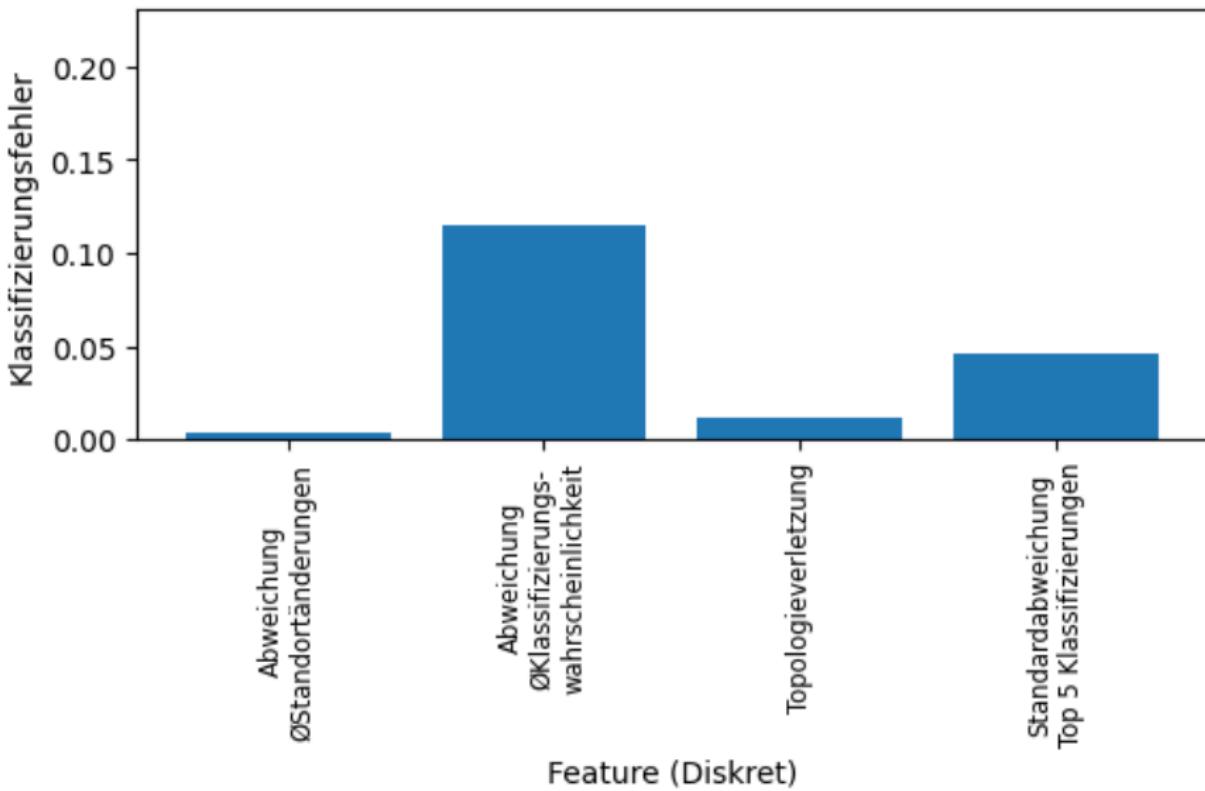
## KLASSIFIZIERUNGSERGEBNISSE I

- FFNN konnte **nicht** erfolgreich trainiert werden
- Entscheidungswald kann 52,52% der Anomalien erkennen
- 2,95% Falsch-Positiv Rate
- Klassifizierungsgenauigkeit abhängig von Erkennungsrate von ML-Modell und Standortkomplexität

## KLASSIFIZIERTUNGSERGEBNISSE II



# PERMUTATIONSWICHTIGKEIT



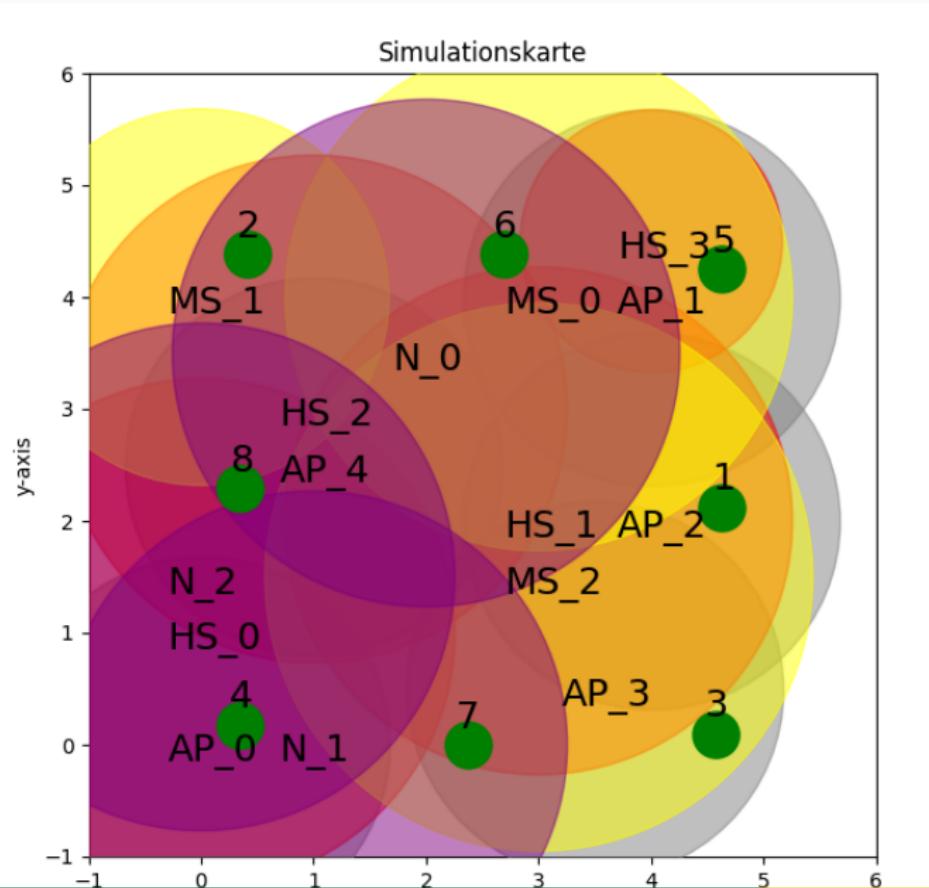
## ERKENNTNISSE DIESER ARBEIT

- Entscheidungswälder skalieren besser als FFNN
- ML-Modelle ohne Rückwärtskante besser als mit
- 98,62% bei 9 Standorten, 87,35% bei 102 Standorten
- Anomalieerkennung: Bis zu 52,58% bei Fehlerrate von 2,95%
- Kleine Datenfenster sind ausreichend
- Geringere Ausführungszeit und klein genug
- Orte zu bestimmen ist besser als Orte und Pfade zu bestimmen
- Feature-Wichtigkeit abhängig vom Einsatzszenario

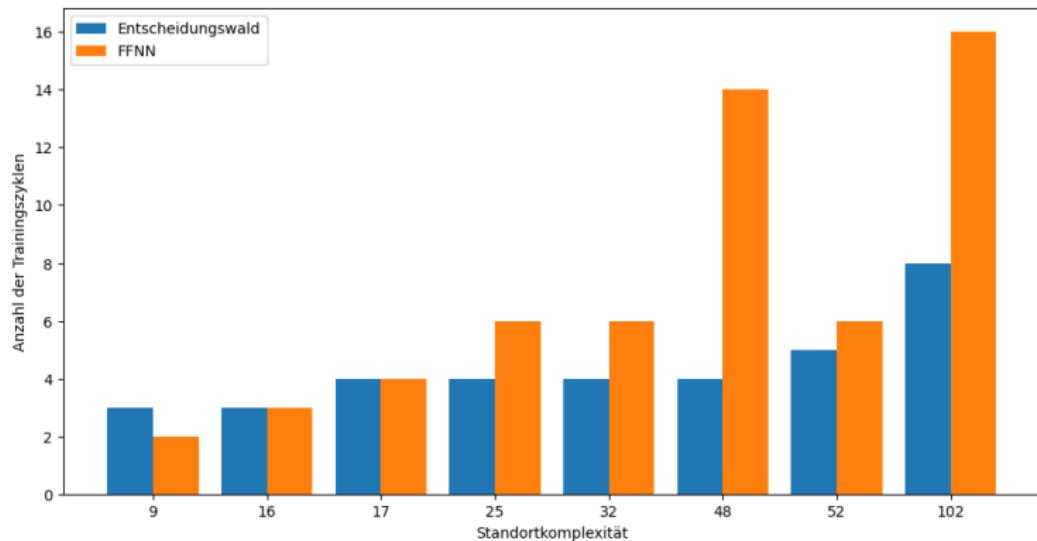
FRAGEN?

DEMO

# SIMULATIONSKARTE VON „SIMPLE SQUARE“



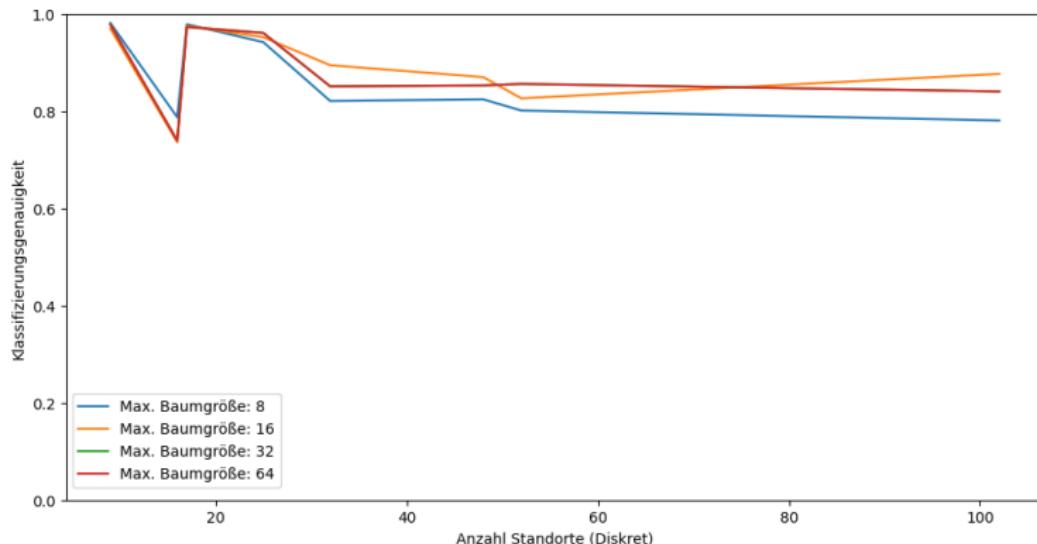
# TRAININGSDATEN PRO ZYKLUS



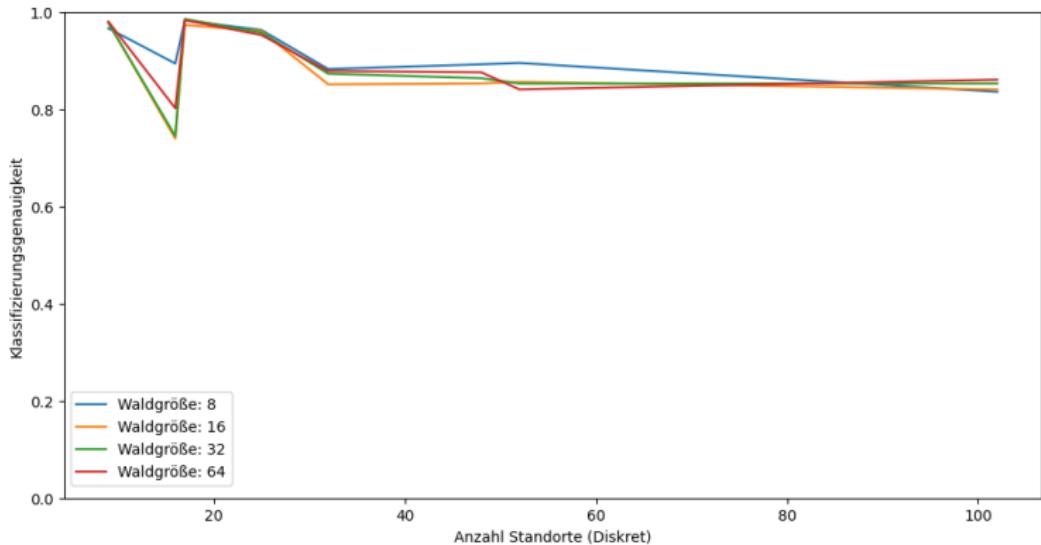
# ANOMALIEERKENNUNG: KLASIFIZIERUNGSERGEBNISSE

| Standorte   | 9       | 16      | 17      | 25      | 32      | 48      | 52      | 102     |
|---|---------|---------|---------|---------|---------|---------|---------|---------|
| <i>P(A)</i>   |         |         |         |         |         |         |         |         |
| DT  | 82,59%  | 81,19%  | 87,14%  | 84,91%  | 79,06%  | 83,47%  | 81,93%  | 76,00%  |
| FFNN  | 77,88%  | 77,88%  | 77,88%  | 77,88%  | 77,88%  | 77,88%  | 77,88%  | 77,88%  |
| Top. (DT)   | 84,77%  | 30,57%  | 83,51%  | 79,76%  | 28,63%  | 24,97%  | 80,55%  | 29,47%  |
| Top. (KNN)  | 86,10%  | 52,17%  | 77,72%  | 79,30%  | 45,06%  | 41,92%  | 74,77%  | 43,55%  |
| Anteil korrekt klassifiziert, indem Anomalie vorlag       |         |         |         |         |         |         |         |         |
| DT  | 34,86%  | 35,52%  | 52,58%  | 50,92%  | 32,21%  | 50,64%  | 23,21%  | 1,92%   |
| FFNN  | 0,00%   | 0,00%   | 0,00%   | 0,00%   | 0,00%   | 0,00%   | 0,00%   | 0,00%   |
| Anteil korrekt klassifiziert, indem keine Anomalie vorlag |         |         |         |         |         |         |         |         |
| DT  | 96,14%  | 94,41%  | 97,05%  | 95,83%  | 92,48%  | 93,13%  | 98,96%  | 97,18%  |
| FFNN  | 100,00% | 100,00% | 100,00% | 100,00% | 100,00% | 100,00% | 100,00% | 100,00% |

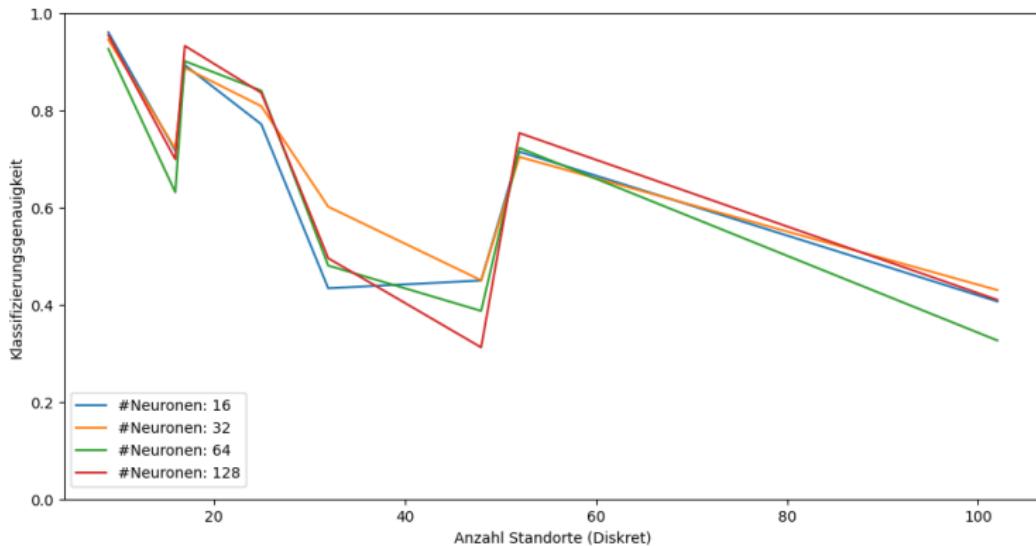
# $P(B \leq 5)_{\text{CONT}}$ - AUSWIRKUNG VON MAXIMALER BAUMHÖHE



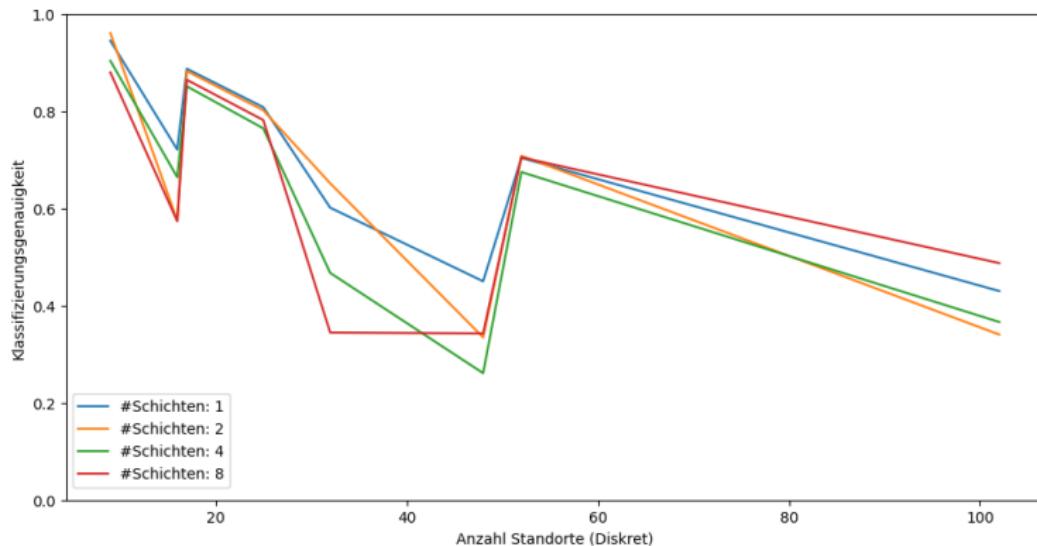
# $P(B \leq 5)_{\text{CONT}}$ - AUSWIRKUNG VON WALDGRÖÙE



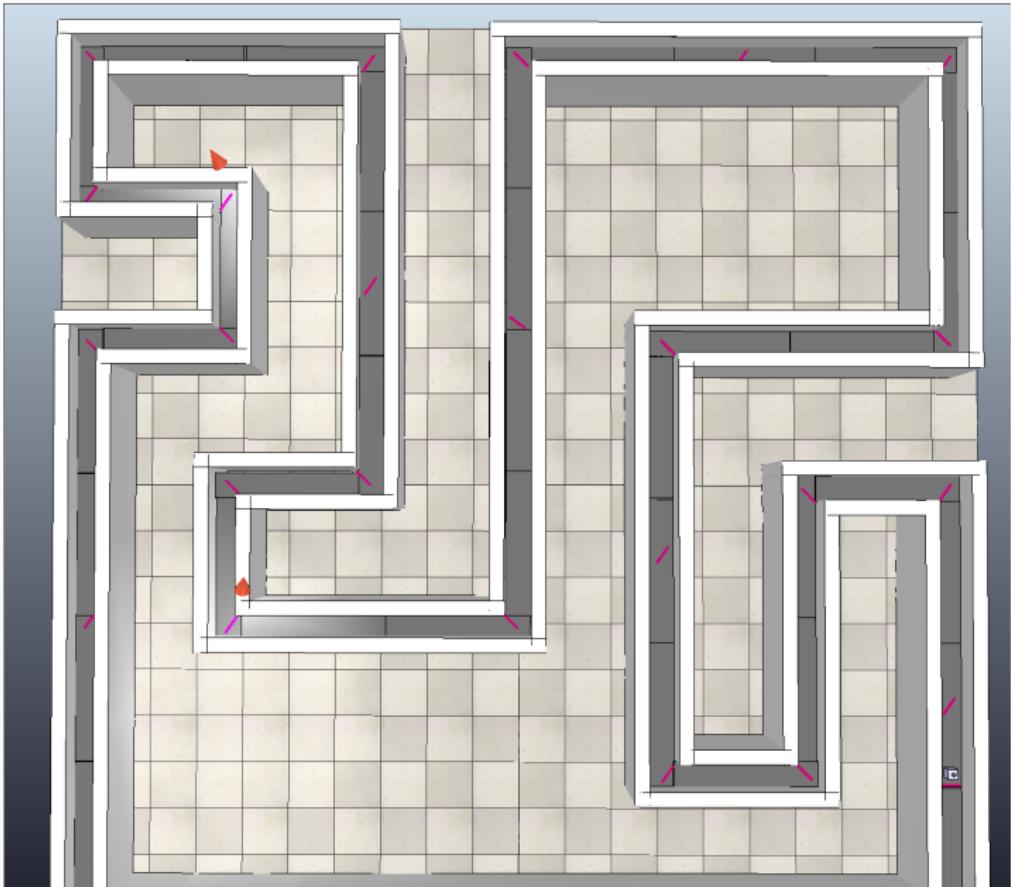
## $P(B \leq 5)_{\text{CONT}}$ - AUSWIRKUNG VON ANZAHL NEURONEN



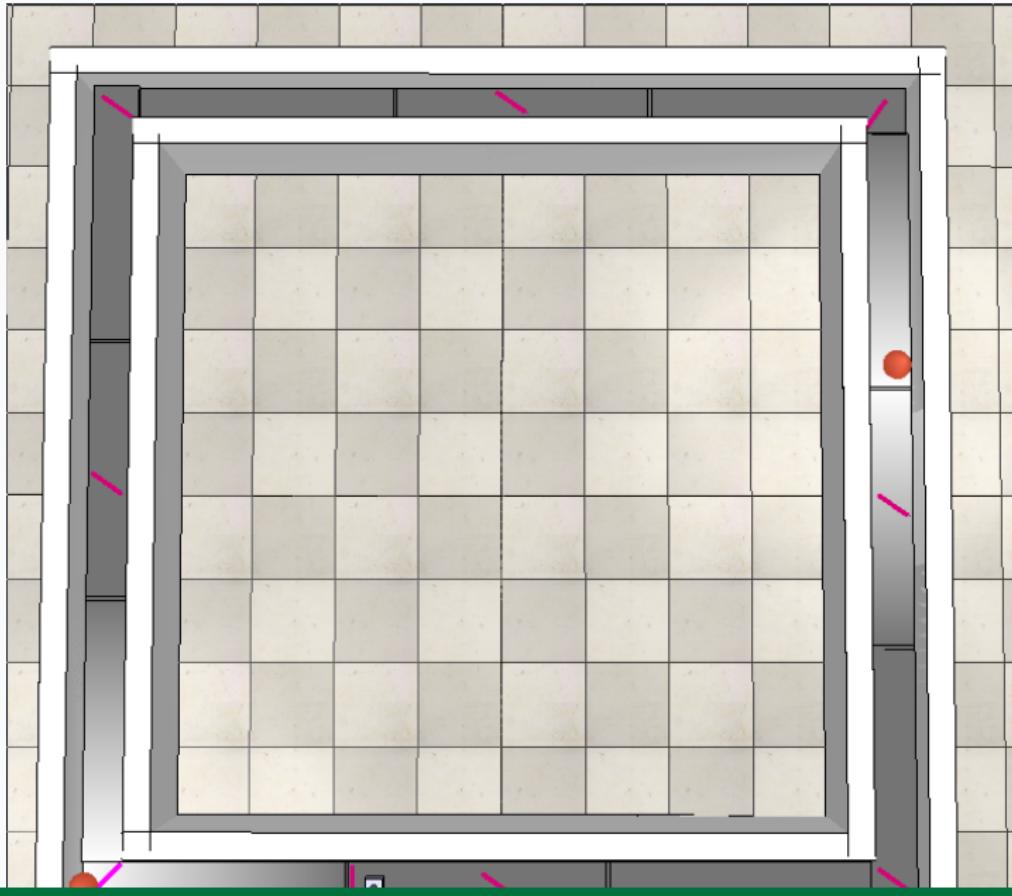
# $P(B \leq 5)_{\text{CONT}}$ - AUSWIRKUNG VON ANZAHL VERDECKTE SCHICHTEN



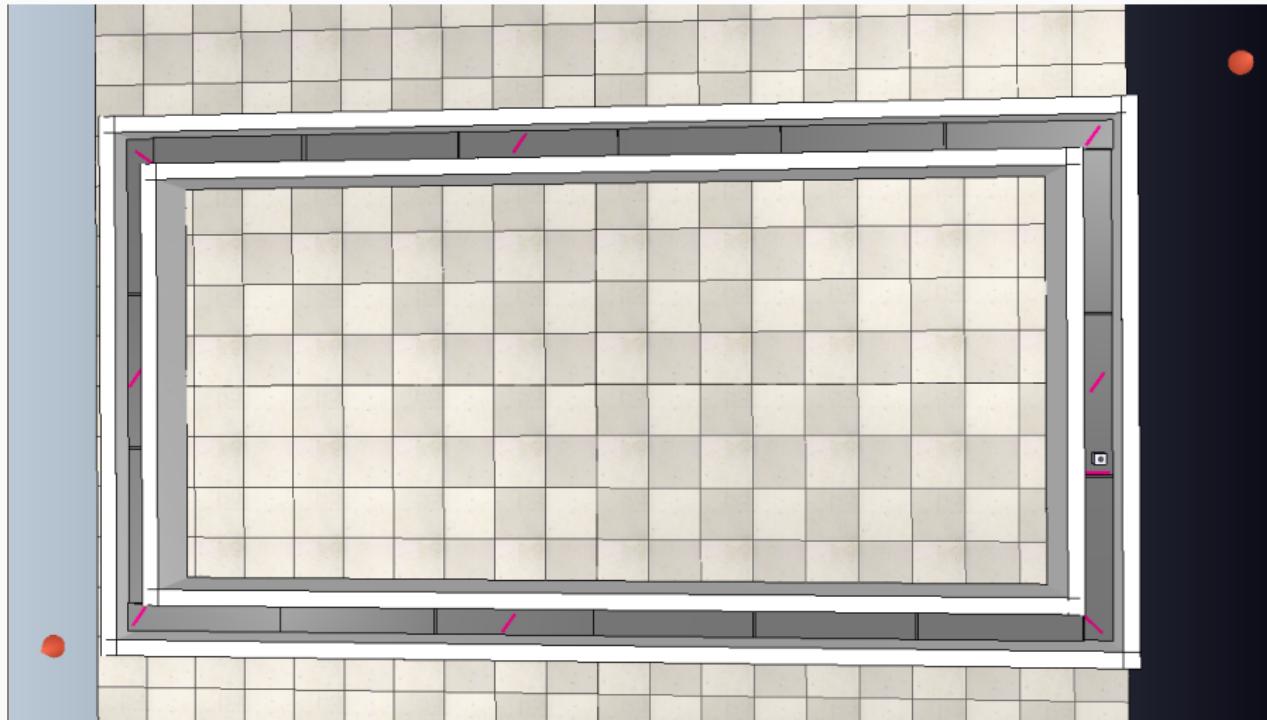
## ROUTE: "MANY CORNERS"



## ROUTE: "SIMPLE SQUARE"



# ROUTE: "LONG RECTANGLE"



## ROUTE: "RECTANGLE WITH RAMP"

