# XAI on Reinforcement Learning: Explaining Tic Tac Toe moves

Data Mining and Machine Learning

Martin Heckel

January 13, 2022

University of Applied Sciences Hof

# Outline

# Introduction

# Introduction

- Tradeoff between performance and explainability of ML models
- Many „real-world“ decisions have to be explainable
- Other decisions have not to be explainable but it helps trusting the model

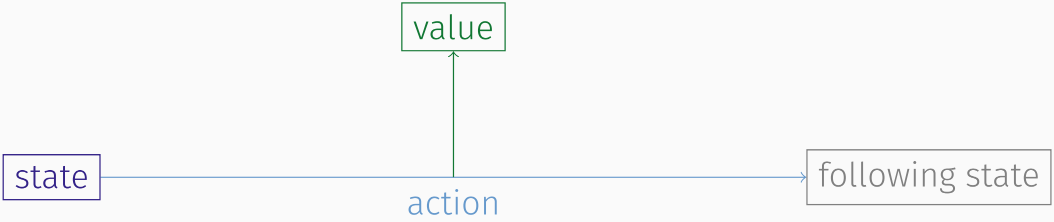
- Trac
- Mar
- Oth

nodel

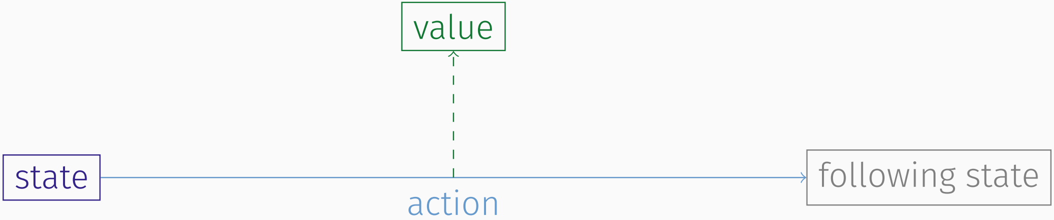
## Machine Learning [1]

# Backgrounds

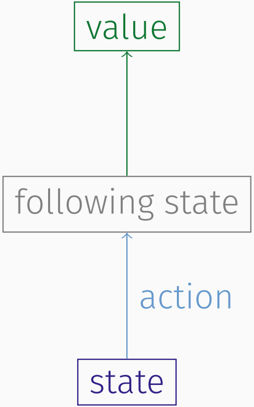
# Reinforcement learning



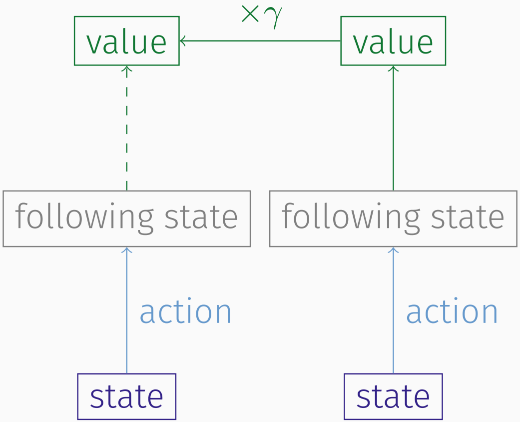
# Reinforcement learning



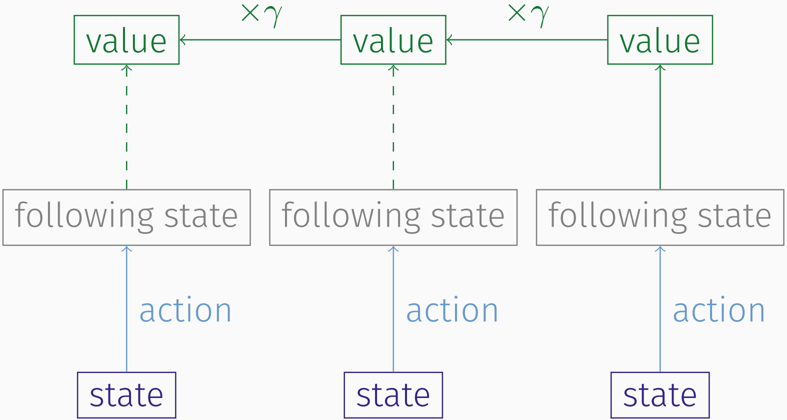
# Reinforcement learning



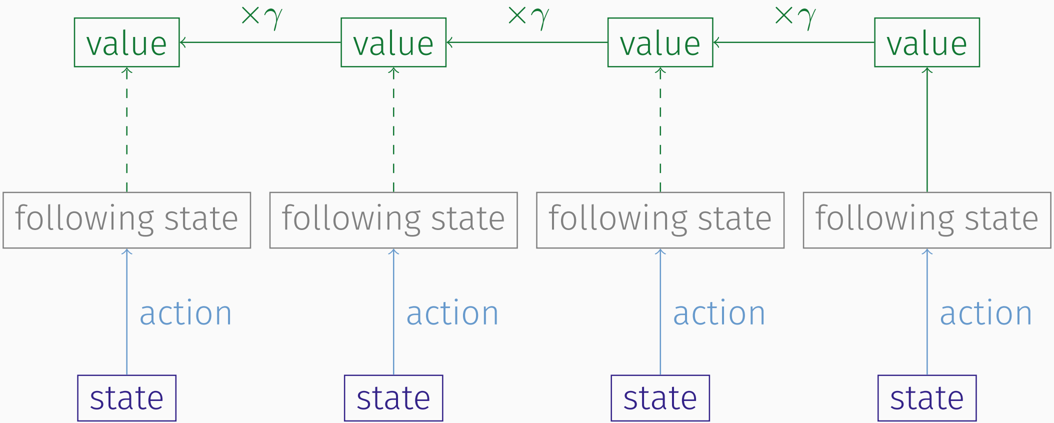
# Reinforcement learning



# Reinforcement learning



# Reinforcement learning



# Tic Tac Toe – Architecture

- **State** is encoded as Vector of size 18
  - „1“ means that a field is taken, „0“ means that it is free
  - First 9 fields represent the game field of the first player
  - Other 9 fields represent the game field of the other player
- **Action** is encoded as a vector of size 9
  - Field with the highest value is used as action
- **Values** are applied in training in a way that the action should have the specified value for the state

# Tic Tac Toe – Architecture

- It is possible to choose between several opponents:
    - **Random** does any random turn (that turn might be invalid)
    - **Easy** does any valid random turn
    - **Medium** does any valid random turn and wins when possible
    - **Hard** does any valid random turns, wins when possible and avoids the opponent to win when possible otherwise
    - **Optimal** implements a backtracking algorithm (minimax) and plays optimal (wins or plays a draw)
    - **Network** uses the trained network to do turns

# eXplainable AI (XAI)

- **Goal**: Explain decisions of a neuronal network in a way that is understandable by humans
- Distinction by the time the explanation is created
  - **Post-Hoc**: Explanation is created after the network has done the prediction
  - **Intrinsic**: Explanation is created while the network does the prediction
- Distinction by relation to the model
  - **Model agnostic**: Independent of the model, generates inputs and uses model outputs for calculation (*black box*)
  - **Model specific**: Dependent of the model, uses its inner workings for calculation (*white box*)
- Distinction by explanation scope
  - **Global**: Explains the model for all possible inputs
  - **Local**: Explains the model for one specific input

# Tic Tac Toe and XAI

- **Hypothesis**: There should be exactly one explanation for a decision made by a deterministic algorithm. Everything else is not an explanation but an interpretation.
- **Initial idea**: Use multiple explanation algorithms to explain the same decision and verify if the return the same explanation

# Tic Tac Toe and XAI

- **Hypothesis**: There should be exactly one explanation for a decision made by a deterministic algorithm. Everything else is not an explanation but an interpretation.
- **Initial idea**: Use multiple explanation algorithms to explain the same decision and verify if the return the same explanation
- **Problem**: Unable to find any library that works with the network layout used (the models I found in the papers were model specific for RL)

# Tic Tac Toe and XAI

- **Hypothesis**: There should be exactly one explanation for a decision made by a deterministic algorithm. Everything else is not an explanation but an interpretation.
- **Initial idea**: Use multiple explanation algorithms to explain the same decision and verify if the return the same explanation
- **Problem**: Unable to find any library that works with the network layout used (the models I found in the papers were model specific for RL)
- **Workaround**: Use a modified version of LIME and manually implement it

# Local Interpretable Model-Agnostic Explanations (LIME)

# Local Interpretable Model-Agnostic Explanations (LIME)

- Can be used to identify „relevant“ parts of images
- Split the image into parts and conduct the prediction for the single parts („remove“ other parts)
- If the predicted probability for the class is higher than for the entire image, that part can be marked as „relevant“
- Explaination is the combination of all „relevant“ parts

# Combining LIME with Tic Tac Toe

- The state of the game is used in input
- The output is a list of action probabilities
- **Idea**: Generate „other versions“ of the sate and predict the action
  - Set the fields that are used by the player or the opponent to „empty“
  - Query the prediction for each generated „other version“ of the state
  - add the predicted value of the resulting state to the sums of all state fields that were active (weight with the distance to the original input)
  - Normalize the results

# Combining LIME with Tic Tac Toe

| | | |
|---|---|---|
| X | | |
| | O | |
| X | X | O |

# Combining LIME with Tic Tac Toe

| | | |
|---|---|---|
| X | | |
| | O | |
| X | X | O |

$$s_0 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}$$

# Combining LIME with Tic Tac Toe

| | | |
|---|---|---|
| X | | |
| | O | |
| X | X | |

$$s_0 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}$$

# Combining LIME with Tic Tac Toe

| | | |
|---|---|---|
| X | | |
| | O | |
| X | X | |

$s_0 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}$

$s_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}$

# Combining LIME with Tic Tac Toe

| | | |
|---|---|---|
| X | | |
| | | |
| X | X | O |

$$s_0 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$s_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

# Combining LIME with Tic Tac Toe

| | | |
|---|---|---|
| X | | |
| | | |
| X | X | O |

$$s_0 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$s_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$s_2 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

# Combining LIME with Tic Tac Toe

| | | |
|---|---|---|
| X | | |
| | | |
| X | X | |

$$s_0 = \begin{array}{|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|}\hline 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ \hline\end{array}$$

$$s_1 = \begin{array}{|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|}\hline 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ \hline\end{array}$$

$$s_2 = \begin{array}{|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|}\hline 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ \hline\end{array}$$

# Combining LIME with Tic Tac Toe

| | | |
|---|---|---|
| X | | |
| | | |
| X | X | |

| | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $s_0 =$ | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 |
| $s_1 =$ | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| $s_2 =$ | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| $s_3 =$ | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

# Combining LIME with Tic Tac Toe

$s_0 = $

| | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 |

# Combining LIME with Tic Tac Toe

| | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $s_0 =$ | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| $a_0 =$ | 0.1 | 0.05 | 0.05 | 0.3 | 0.1 | 0.2 | 0.05 | 0.05 | 0.1 |

# Combining LIME with Tic Tac Toe

$s_0 =$

| | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 |

$a_0 =$

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 0.1 | 0.05 | 0.05 | 0.3 | 0.1 | 0.2 | 0.05 | 0.05 | 0.1 |

$s_1 =$

| | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |

$a_1 =$

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 0.01 | 0.05 | 0.0 | 0.6 | 0.9 | 0.05 | 0.05 | 0.05 | 0.1 |

# Combining LIME with Tic Tac Toe

$$s_0 = \begin{array}{|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|} \hline 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ \hline \end{array}$$

$$a_0 = \begin{array}{|c|c|c|c|c|c|c|c|c|} \hline 0.1 & 0.05 & 0.05 & 0.3 & 0.1 & 0.2 & 0.05 & 0.05 & 0.1 \\ \hline \end{array}$$

$$s_1 = \begin{array}{|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|} \hline 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ \hline \end{array}$$

$$a_1 = \begin{array}{|c|c|c|c|c|c|c|c|c|} \hline 0.01 & 0.05 & 0.0 & 0.6 & 0.9 & 0.05 & 0.05 & 0.05 & 0.1 \\ \hline \end{array}$$

$$s_2 = \begin{array}{|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|} \hline 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ \hline \end{array}$$

$$a_2 = \begin{array}{|c|c|c|c|c|c|c|c|c|} \hline 0.1 & 0.1 & 0.05 & 0.25 & 0.1 & 0.2 & 0.05 & 0.05 & 0.1 \\ \hline \end{array}$$

# Combining LIME with Tic Tac Toe

$$s_0 = \begin{array}{|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|} \hline 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ \hline \end{array}$$

$$a_0 = \begin{array}{|c|c|c|c|c|c|c|c|c|} \hline 0.1 & 0.05 & 0.05 & 0.3 & 0.1 & 0.2 & 0.05 & 0.05 & 0.1 \\ \hline \end{array}$$

$$s_1 = \begin{array}{|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|} \hline 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ \hline \end{array}$$

$$a_1 = \begin{array}{|c|c|c|c|c|c|c|c|c|} \hline 0.01 & 0.05 & 0.0 & 0.6 & 0.9 & 0.05 & 0.05 & 0.05 & 0.1 \\ \hline \end{array}$$

$$s_2 = \begin{array}{|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|} \hline 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ \hline \end{array}$$

$$a_2 = \begin{array}{|c|c|c|c|c|c|c|c|c|} \hline 0.1 & 0.1 & 0.05 & 0.25 & 0.1 & 0.2 & 0.05 & 0.05 & 0.1 \\ \hline \end{array}$$

$$s_3 = \begin{array}{|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|} \hline 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline \end{array}$$

$$a_3 = \begin{array}{|c|c|c|c|c|c|c|c|c|} \hline 0.1 & 0.0 & 0.05 & 0.4 & 0.1 & 0.2 & 0.0 & 0.05 & 0.1 \\ \hline \end{array}$$

# Combining LIME with Tic Tac Toe

$$s_0 = \begin{array}{|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|}\hline 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ \hline\end{array}$$

$$a_0 = \begin{array}{|c|c|c|c|c|c|c|c|c|}\hline 0.1 & 0.05 & 0.05 & \boxed{0.3} & 0.1 & 0.2 & 0.05 & 0.05 & 0.1 \\ \hline\end{array}$$

$$s_1 = \begin{array}{|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|}\hline 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ \hline\end{array}$$

$$a_1 = \begin{array}{|c|c|c|c|c|c|c|c|c|}\hline 0.01 & 0.05 & 0.0 & \boxed{0.6} & 0.9 & 0.05 & 0.05 & 0.05 & 0.1 \\ \hline\end{array}$$

$$s_2 = \begin{array}{|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|}\hline 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ \hline\end{array}$$

$$a_2 = \begin{array}{|c|c|c|c|c|c|c|c|c|}\hline 0.1 & 0.1 & 0.05 & \boxed{0.25} & 0.1 & 0.2 & 0.05 & 0.05 & 0.1 \\ \hline\end{array}$$

$$s_3 = \begin{array}{|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|}\hline 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline\end{array}$$

$$a_3 = \begin{array}{|c|c|c|c|c|c|c|c|c|}\hline 0.1 & 0.0 & 0.05 & \boxed{0.4} & 0.1 & 0.2 & 0.0 & 0.05 & 0.1 \\ \hline\end{array}$$

# Combining LIME with Tic Tac Toe

$$s_0 = \begin{array}{|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|} \hline 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ \hline \end{array}$$

$$a_0 = \begin{array}{|c|c|c|c|c|c|c|c|c|} \hline 0.1 & 0.05 & 0.05 & \boxed{0.3} & 0.1 & 0.2 & 0.05 & 0.05 & 0.1 \\ \hline \end{array}$$

# Combining LIME with Tic Tac Toe

$s_0 =$

| | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 |

$a_0 =$

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 0.1 | 0.05 | 0.05 | 0.3 | 0.1 | 0.2 | 0.05 | 0.05 | 0.1 |

$v_s =$

| | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

# Combining LIME with Tic Tac Toe

$$s_0 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$a_0 = \begin{bmatrix} 0.1 & 0.05 & 0.05 & 0.3 & 0.1 & 0.2 & 0.05 & 0.05 & 0.1 \end{bmatrix}$$

$$v_s = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

# Combining LIME with Tic Tac Toe

$$s_0 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$a_0 = \begin{bmatrix} 0.1 & 0.05 & 0.05 & 0.3 & 0.1 & 0.2 & 0.05 & 0.05 & 0.1 \end{bmatrix}$$

$$v_s = \begin{bmatrix} 0.3 & 0 & 0 & 0 & 0 & 0 & 0.3 & 0.3 & 0 & 0 & 0 & 0 & 0 & 0.3 & 0 & 0 & 0 & 0.3 \end{bmatrix}$$

# Combining LIME with Tic Tac Toe

$s_0 =$

| | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 |

$a_0 =$

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 0.1 | 0.05 | 0.05 | 0.3 | 0.1 | 0.2 | 0.05 | 0.05 | 0.1 |

$s_1 =$

| | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |

$a_1 =$

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 0.1 | 0.01 | 0.0 | $\boxed{0.6}$ | 0.04 | 0.09 | 0.05 | 0.05 | 0.06 |

$v_s =$

| | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.3 | 0 | 0 | 0 | 0 | 0 | 0.3 | 0.3 | 0 | 0 | 0 | 0 | 0 | 0.3 | 0 | 0 | 0 | 0.3 |

$\times \frac{2}{3}$

# Combining LIME with Tic Tac Toe

$$s_0 = \begin{array}{|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|}\hline 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ \hline \end{array}$$

$$a_0 = \begin{array}{|c|c|c|c|c|c|c|c|c|}\hline 0.1 & 0.05 & 0.05 & 0.3 & 0.1 & 0.2 & 0.05 & 0.05 & 0.1 \\ \hline \end{array}$$

$$s_1 = \begin{array}{|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|}\hline 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ \hline \end{array}$$

$$a_1 = \begin{array}{|c|c|c|c|c|c|c|c|c|}\hline 0.1 & 0.01 & 0.0 & \boxed{0.6} & 0.04 & 0.09 & 0.05 & 0.05 & 0.06 \\ \hline \end{array}$$

$$v_s = \begin{array}{|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|}\hline 0.3 & 0 & 0 & 0 & 0 & 0 & 0.3 & 0.3 & 0 & 0 & 0 & 0 & 0 & 0.3 & 0 & 0 & 0 & 0.3 \\ \hline \end{array}$$

$$\times \frac{2}{3}$$

# Combining LIME with Tic Tac Toe

$$
\begin{array}{l}
s_0 = [\,1 \mid 0 \mid 0 \mid 0 \mid 0 \mid 0 \mid 1 \mid 1 \mid 0 \mid 0 \mid 0 \mid 0 \mid 0 \mid 1 \mid 0 \mid 0 \mid 0 \mid 1\,] \\
a_0 = [\,0.1 \mid 0.05 \mid 0.05 \mid 0.3 \mid 0.1 \mid 0.2 \mid 0.05 \mid 0.05 \mid 0.1\,] \\
s_1 = [\,1 \mid 0 \mid 0 \mid 0 \mid 0 \mid 0 \mid 1 \mid 1 \mid 0 \mid 0 \mid 0 \mid 0 \mid 0 \mid 1 \mid 0 \mid 0 \mid 0 \mid 0\,] \\
a_1 = [\,0.1 \mid 0.01 \mid 0.0 \mid \boxed{0.6} \mid 0.04 \mid 0.09 \mid 0.05 \mid 0.05 \mid 0.06\,] \\
v_s = [\,0.7 \mid 0 \mid 0 \mid 0 \mid 0 \mid 0 \mid 0.7 \mid 0.7 \mid 0 \mid 0 \mid 0 \mid 0 \mid 0 \mid 0.7 \mid 0 \mid 0 \mid 0 \mid 0.3\,]
\end{array}
$$

# Combining LIME with Tic Tac Toe

| | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $s_0 =$ | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| $a_0 =$ | 0.1 | 0.05 | 0.05 | 0.3 | 0.1 | 0.2 | 0.05 | 0.05 | 0.1 |

| | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $s_1 =$ | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| $a_1 =$ | 0.1 | 0.01 | 0.0 | 0.6 | 0.04 | 0.09 | 0.05 | 0.05 | 0.06 |

| | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $v_s =$ | 0.7 | 0 | 0 | 0 | 0 | 0 | 0.7 | 0.7 | 0 | 0 | 0 | 0 | 0 | 0.7 | 0 | 0 | 0 | 0.3 |

# Combining LIME with Tic Tac Toe

$s_0 =$

| | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 |

$a_0 =$

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 0.1 | 0.05 | 0.05 | 0.3 | 0.1 | 0.2 | 0.05 | 0.05 | 0.1 |

$s_1 =$

| | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |

$a_1 =$

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 0.1 | 0.01 | 0.0 | 0.6 | 0.04 | 0.09 | 0.05 | 0.05 | 0.06 |

$v_s =$

| | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.7 | 0 | 0 | 0 | 0 | 0 | 0.7 | 0.7 | 0 | 0 | 0 | 0 | 0 | 0.7 | 0 | 0 | 0 | 0.3 |

$v_n =$

| | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.23 | 0 | 0 | 0 | 0 | 0 | 0.23 | 0.23 | 0 | 0 | 0 | 0 | 0 | 0.23 | 0 | 0 | 0 | 0.08 |

# Combining LIME with Tic Tac Toe

$$v_n = \begin{array}{|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|}\hline 0.23 & 0 & 0 & 0 & 0 & 0 & 0.23 & 0.23 & 0 & 0 & 0 & 0 & 0 & 0.23 & 0 & 0 & 0 & 0.08 \\ \hline\end{array}$$

# Combining LIME with Tic Tac Toe

$v_n =$

| | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.23 | 0 | 0 | 0 | 0 | 0 | 0.23 | 0.23 | 0 | 0 | 0 | 0 | 0 | 0.23 | 0 | 0 | 0 | 0.08 |

$v =$

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

# Combining LIME with Tic Tac Toe

$$v_n = \begin{array}{|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|} \hline 0.23 & 0 & 0 & 0 & 0 & 0 & 0.23 & 0.23 & 0 & 0 & 0 & 0 & 0 & 0.23 & 0 & 0 & 0 & 0.08 \\ \hline \end{array}$$

$$v = \begin{array}{|c|c|c|c|c|c|c|c|c|} \hline 0.23 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline \end{array}$$

# Combining LIME with Tic Tac Toe

$$v_n = \begin{array}{|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|} \hline 0.23 & 0 & 0 & 0 & 0 & 0 & 0.23 & 0.23 & 0 & 0 & 0 & 0 & 0 & 0.23 & 0 & 0 & 0 & 0.08 \\ \hline \end{array}$$

$$v = \begin{array}{|c|c|c|c|c|c|c|c|c|} \hline 0.23 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline \end{array}$$

# Combining LIME with Tic Tac Toe

$$v_n = \begin{bmatrix} 0.23 & 0 & 0 & 0 & 0 & 0 & 0.23 & 0.23 & 0 & 0 & 0 & 0 & 0 & 0.23 & 0 & 0 & 0 & 0.08 \end{bmatrix}$$

$$v = \begin{bmatrix} 0.23 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

# Combining LIME with Tic Tac Toe

$v_n =$

| | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.23 | 0 | 0 | 0 | 0 | 0 | 0.23 | 0.23 | 0 | 0 | 0 | 0 | 0 | 0.23 | 0 | 0 | 0 | 0.08 |

$v =$

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 0.23 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

# Combining LIME with Tic Tac Toe

$v_n =$

| | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.23 | 0 | 0 | 0 | 0 | 0 | 0.23 | 0.23 | 0 | 0 | 0 | 0 | 0 | 0.23 | 0 | 0 | 0 | 0.08 |

$v =$

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 0.23 | 0 | 0 | 0 | 0.23 | 0 | 0 | 0 | 0 |

# Combining LIME with Tic Tac Toe

$$v_n = \begin{bmatrix} 0.23 & 0 & 0 & 0 & 0 & 0 & 0.23 & 0.23 & 0 & 0 & 0 & 0 & 0 & 0.23 & 0 & 0 & 0 & 0.08 \end{bmatrix}$$

$$v = \begin{bmatrix} 0.23 & 0 & 0 & 0 & 0.23 & 0 & 0 & 0 & 0 \end{bmatrix}$$

# Combining LIME with Tic Tac Toe

$$v_n = \begin{array}{|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|} \hline 0.23 & 0 & 0 & 0 & 0 & 0 & 0.23 & 0.23 & 0 & 0 & 0 & 0 & 0 & 0.23 & 0 & 0 & 0 & 0.08 \\ \hline \end{array}$$

$$v = \begin{array}{|c|c|c|c|c|c|c|c|c|} \hline 0.23 & 0 & 0 & 0 & 0.23 & 0 & 0.23 & 0 & 0 \\ \hline \end{array}$$

# Combining LIME with Tic Tac Toe

$v_n =$

| | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.23 | 0 | 0 | 0 | 0 | 0 | 0.23 | 0.23 | 0 | 0 | 0 | 0 | 0 | 0.23 | 0 | 0 | 0 | 0.08 |

$v =$

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 0.23 | 0 | 0 | 0 | 0.23 | 0 | 0.23 | 0.23 | 0 |

# Combining LIME with Tic Tac Toe

$v_n =$

| | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.23 | 0 | 0 | 0 | 0 | 0 | 0.23 | 0.23 | 0 | 0 | 0 | 0 | 0 | 0.23 | 0 | 0 | 0 | 0.08 |

$v =$

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 0.23 | 0 | 0 | 0 | 0.23 | 0 | 0.23 | 0.23 | 0.08 |

# Combining LIME with Tic Tac Toe

$$v_n = \begin{bmatrix} 0.23 & 0 & 0 & 0 & 0 & 0 & 0.23 & 0.23 & 0 & 0 & 0 & 0 & 0 & 0.23 & 0 & 0 & 0 & 0.08 \end{bmatrix}$$

$$v = \begin{bmatrix} 0.23 & 0 & 0 & 0 & 0.23 & 0 & 0.23 & 0.23 & 0.08 \end{bmatrix}$$

| | | |
|---|---|---|
| X | | |
| | O | |
| X | X | O |

# Demo 1

Does this work at all?

# Demo 2

Inspect the ML model

# Questions?

[1] Randall Munroe. *Machine Learning*. URL: https://xkcd.com/1838/ (besucht am 11. 01. 2022).