General points, that could not meet the scientific rigor:

- BOLD model is constructed through the alignment levels. Am I giving the model too much information through the BOLD model, since I want to estimate the next alignment level? Or, even more pressing: Am I giving, in the case of H2, the model actually the correct alignment level?

  >> For H3: Check, if the CNN model places its weights primarily on the BOLD model, which would imply that the main source of the prediction are the previous alignment levels that may be extracted from the modeled BOLD signal (see: checking the weights).

  Weights of CNN filters for the measurements and for the

  >> For H2: Check: How good would classification work, if model would have not been given? This result would give insights into how much the model relied on the BOLD signal.

  >> For H2: Check:

  My baseline may be unfair. The baseline model currently always predicts the average alignment level. To have an equal comparison, it might be fair to let it predict the last alignment level? (See: Changing the baseline model)

For H2 (the real-time estimation):

Checking the weights:

- Concern:
  The CNN model solely relies on the BOLD model – and not on the measurements.
- Approach:
  Look at weights of the different layers of the model (convolutional layer; the dense layer after the convolutional layer and a pooling layer) and check, if the weights corresponding to the BOLD layer are comparably high in regards to the other weights.
- Results:
  Weights of convolutional layer for measurements:

  Weights of convolutional layer for the BOLD signal:

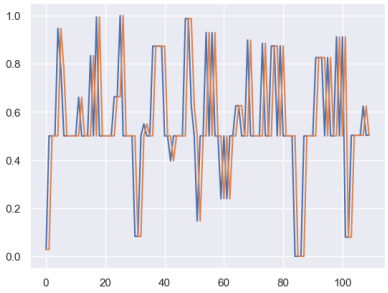
  Weights of the dense layer:

  
- Conclusion:
  The weights that are associated with the BOLD model are not significantly higher than the other weights – thus, the model infers its results also from the measured data.
- Discussion:

This is certainly no proof that the model infers information correctly from the measured data – it could, for example, be that the model solely infers “noise” from the measured signals. This is, however, that leads to the measured MSE and MAE errors, while not

Changing the baseline model:

- Concern: In the report I had included a baseline model, that always predicted the average alignment level. This may be unfair, since the CNN model gets the BOLD signal, that is computed from previous alignment levels.
- Approach: Creating a new baseline model, that always predicts the alignment level of the previous time step (see plot). The alignment level and the prediction from the baseline model have identical values in less than half of the timesteps (in 47.27% of timesteps).



- Result: MSE: 0.0682, MAE: 0.1727