1. Pytorch recap
2. Custom layer
   1. Not fully connected
   2. No biases
   3. Binary weights (if(abs(wegiht) < thresh) then weight 🡺 0, else(weight 🡺 1))
      1. STE method
3. Custom loss function
   1. Penalty & penalty strength

Next steps:

* Before the simulations wrap the entire model definition and the training loop into a function that takes as arguments all hyperparameters aswell as the optimizer (this function should rerun the model under the specified conditions many times to get a distribution of the performance metrics)
* Simulate
  + Don’t change any hyperparameters yet, just try out different optimizers with different options
  + Then simulate for different values of the cutoff parameter and the penalty strength
  + Also look at the architecture after the masking layer