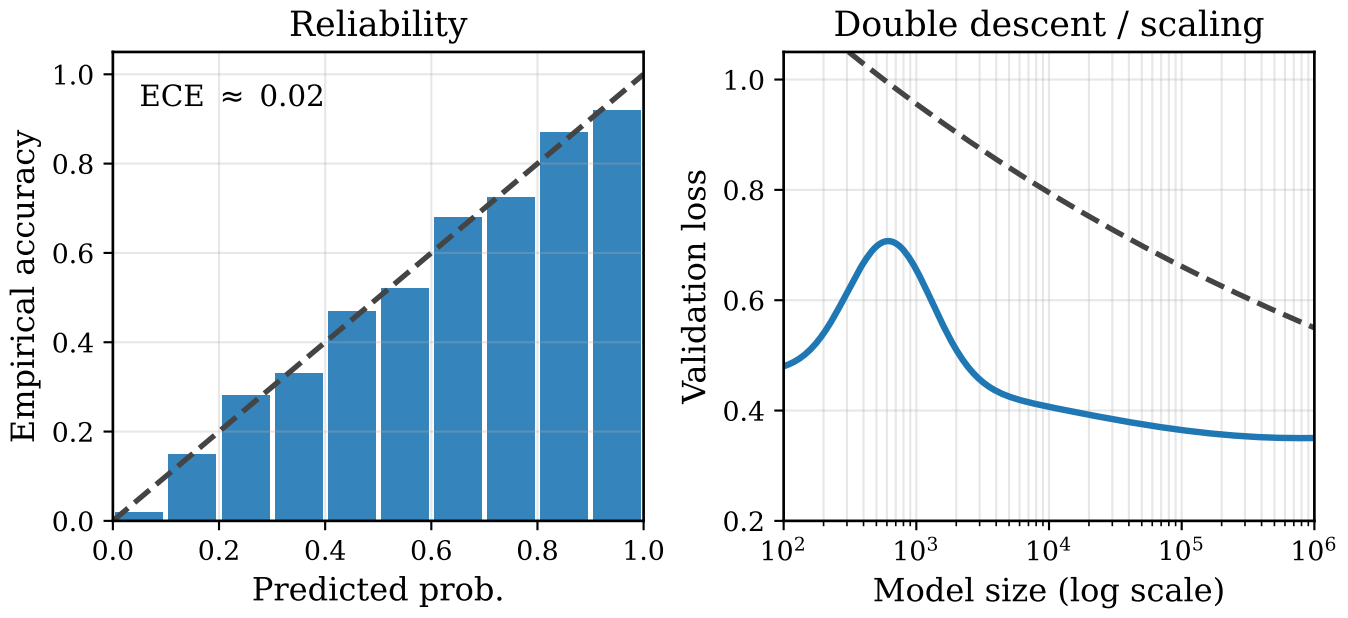
Reliability
ECE ≈ 0.02
Empirical accuracy
Predicted prob.
0.0
0.2
0.4
0.6
0.8
1.0
Double descent / scaling
Validation loss
Model size (log scale)
10^2
10^3
10^4
10^5
10^6