



1. Fit a discrete auto-encoder with learned prior

$$\mathcal{L} = -\log p_W(W) - \log p(Z | f(W))$$

2. Measure complexity terms

$$K(Z) = K(p_W) + K(W | p_W) + K(f) + K(Z | W, f)$$