HW3: Restricted Boltzman Machine

Jens Ifver

September 2021

The Restricted Boltzman Machine (RBM) we use in this task consist of three visible neurons V_j and M hidden neurons h_i where $M = 1, 2, 4, 8$. We use the CD-k algorithm to train the RBM to learn the distribution of the XOR-dataset where four patterns are given 1/4 probability and the rest zero. We use the CD-k algorithm to adjust the weights and thresholds so that the Boltzman distribution $P_B(x)$ approximate the input distribution $P_{data}(x)$. To estimate $P_B(x)$ we iteratively feed random 3-dimensional Boolean patterns and use the weights and thresholds obtained in the CD-k algorithm to apply neuron updates and then keep count of the number of times the state converge to every pattern. In theory $M = 2^N/2 - 1$ hidden neurons are enough for $P_B(x)$ to get very close to $P_{data}(x)$. To measure how similar the two distributions are we use the Kullback-Leibler divergence D_{KL}. In Figure 1 we see that D_{KL} decreases as the number of hidden neurons increase. This is due to the fact that the hidden neurons encode correlations between the visible neurons and each of the hidden neurons can describe one of the binary patterns. Therefor it makes sense that for $M = 1, 2$ the distributions have a higher divergence and for $M = 4, 8$ there is a lower divergence.

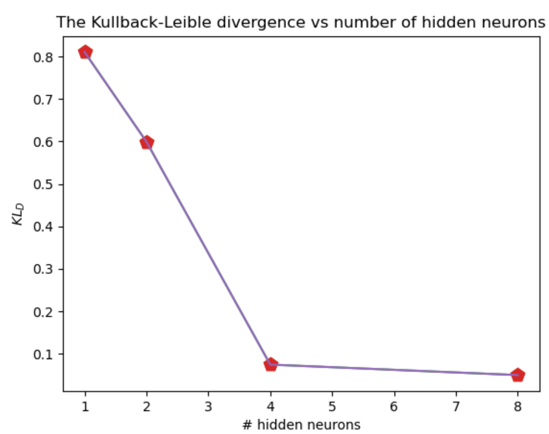

Figure 1: The figure shows the computed Kullback-Leibler divergence as a function of number of hidden neurons M in an RBM. Results were obtained using k=100, $\eta = 0.1$ and 200 trials for each M. The red dots represent the Kullback-Leibler divergence computed for $M = 1, 2, 4, 8$.