

Restricted Boltzmann Machine

Trilokinath Modi

The KL divergence approximates the difference between model distribution and data distribution. To make the machine learn, sampling is required. For each iteration, 3 randomly chosen samples are drawn and fed to the machine for weight and thresholds update. To avoid numerical inconsistency due to $\log(\infty)$ at the beginning, the KL divergence is set to a large value, here 10, for each pattern in data distribution resulting 140 as initial value for KL divergence. Boltzmann probability is computed for all patterns after every 100 sets of samples drawn i.e. 300 samples.

To determine how long the dynamics must run, the threshold to stop iterations is kept as 0.1 i.e. when net KL divergence is < 0.1 , the training stop. However, the upper limit of maximum possible iterations is tested up to $1e5$ and it deduced that $M = 2,4$ doesn't converge even at $1e5$ iterations, whereas $M = 8,16$ needs roughly $< 40000, < 10000$ iterations respectively for stopping criteria. Figure 1(a) reflects the number of iterations required for $M \in \{2,4,8,16\}$ to reach stopping criteria¹. At steady state, incomplete patterns is fed to the machine for 10 iterations of McCulloh-Pitts dynamics. The dynamics is shown in fig 1(b)².

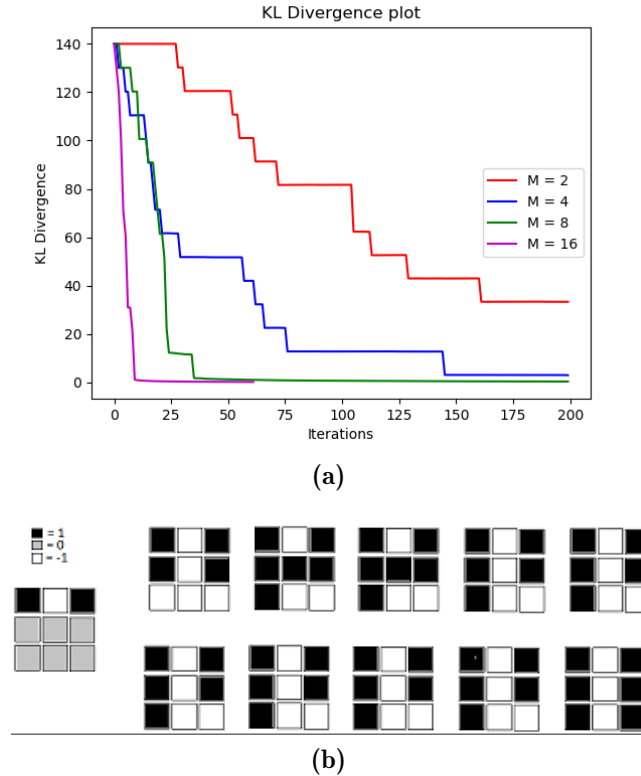


Figure 1: (a) portrays the convergence of algorithm for different number of neurons. For $M = 2,4$, the algorithm doesn't meet stopping criteria of $D < 0.1$, whereas for $M = 8,16$, the algorithm reaches stopping criteria at 2,2 respectively. The x-axis is divided by 100 for aesthetic.(b) Indicate the 10 McCulloh pitts dynamics ran on trained machine with 16 neuron and an incomplete pattern was fed. The algorithm converged this incomplete pattern to expected pattern at 10th McCulloh iteration.

¹Note that for this figure, the max iterations was reduced to 20000, however, 100000 is also tested

²Many other incomplete patterns were tested and often they converged to required pattern at first McCulloh pitt dynamic.