<u>Restricted Boltzmann machine</u>

The program written for RBM trained is using the CD-k algorithm. 1000 random patterns are used and the CD-K algorithm iterates with a probability of 0.25 with k = 200 and for 20 mini batches. Also, the learning set is set to 0.005. The trained machine is provided with 3000 random patterns to compute Dkl, and each pattern's dynamics were iterated for 2000 cycles. The output was compared to the patterns for each round, i.e., 6*10^6 times (3000*2000). The data distribution of the RBM(PB) was used to compute the probability that each pattern would occur. To determine the Dkl, PB is then compared with the input data distribution PD.

Formula for the upper bound of Dkl:

$$D_{KL} \leq \begin{cases} log2[N - \lfloor log_2(M+1) \rfloor - \frac{M+1}{2^{log_2(M+1)}}], & M < 2^{(N-1)} - 1 \\ 0, & M \geq 2^{(N-1)} - 1 \end{cases}$$

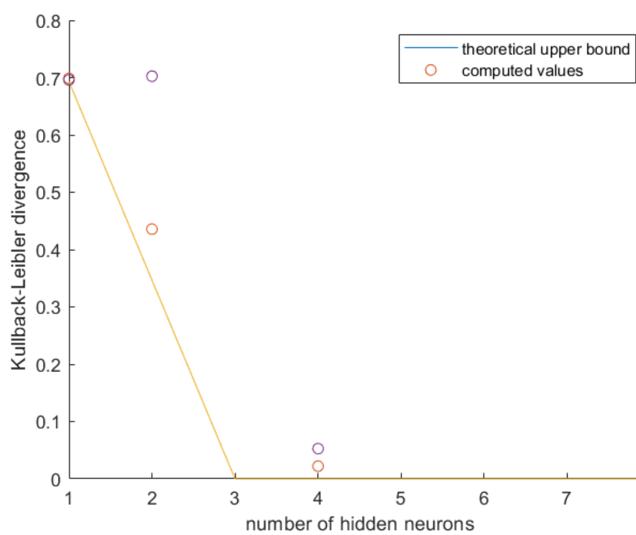

Figure 1: The graph highlights the numerical estimates of the Kullback-Leibler divergence against the number of hidden neurons, in comparison with the upper bound(theoretical).

We can conclude from the graph that as the Dkl approaches 0 the data distribution of the RBM approaches the input data distribution. However, it is possible that the experimental Dkl occasionally exceeds the upper bound, therefore it is not necessary that the CD-k method will always converge to the ideal solution. An important observation is a larger learning rate will result in the failure of convergence of the algorithm.