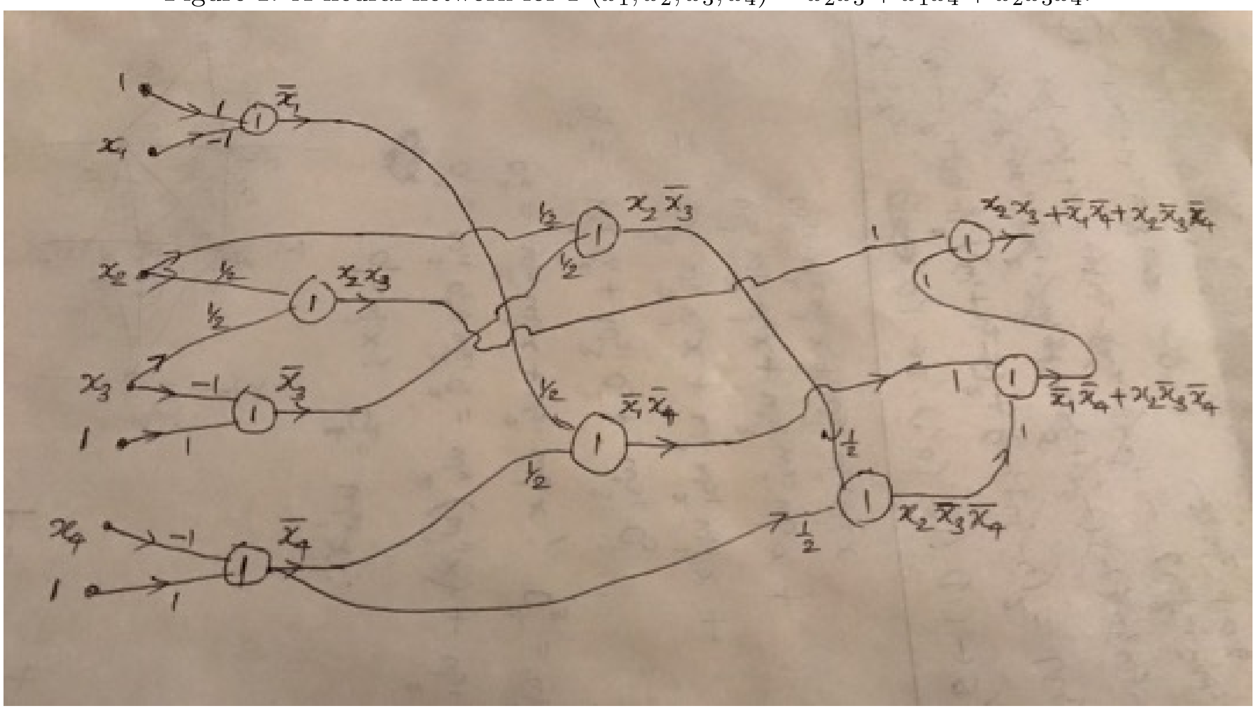
CSE 4502/5717 Big Data Analytics. Fall 2019

Exam III (model) Solutions

1. The loss function is $L(w_1, w_2) = (w_2 - 2)^2 + (w_1 - 1)^2 + (w_1 + w_2 - 5)^2 + (2w_1 + w_2 - 4)^2$ $= 6w_1^2 + 3w_2^2 + 6w_1w_2 - 28w_1 - 22w_2 + 46$. We want to have: $\frac{\partial L}{\partial w_1} = 0$ and $\frac{\partial L}{\partial w_2} = 0$.
   $\frac{\partial L}{\partial w_1} = 0$ implies that $12w_1 + 6w_2 = 28$ and $\frac{\partial L}{\partial w_2} = 0$ implies that $6w_1 + 6w_2 = 22$. Solving these two equations, we get: $w_1 = 1$ and $w_2 = \frac{8}{3}$.

2. Here is a multilevel perceptron for realizing the Boolean function $F(x_1, x_2, x_3, x_4) = x_2x_3 + \bar{x_1}\bar{x_4} + x_2\bar{x_3}\bar{x_4}$:

Figure 1: A neural network for $F(x_1, x_2, x_3, x_4) = x_2x_3 + \bar{x_1}\bar{x_4} + x_2\bar{x_3}\bar{x_4}$.



3. Let

$$\boldsymbol{X} = \begin{bmatrix} x_1^1 & x_1^2 & \cdots & x_1^n \\ x_2^1 & x_2^2 & \cdots & x_2^n \\ \cdots & & & \\ x_m^1 & x_m^2 & \cdots & x_m^n \end{bmatrix}.$$

   Then, we showed that the optimal value for $\boldsymbol{w}$ is $(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{y}$ where $\boldsymbol{y} = (y_1\ y_2\ \cdots\ y_m)^T$.
   We can compute $\boldsymbol{X}^T\boldsymbol{X}$ in $O(\log m)$ time using $\frac{n^2m}{\log m}$ processors. Followed by this, we can compute $(\boldsymbol{X}^T\boldsymbol{X})^{-1}$ in $O(\log^2 n)$ time using n^4 processors. Next, we can compute $(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T$

in $O(\log n)$ time using $\frac{n^2 m}{\log n}$ processors. Finally, we can compute $(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{y}$ in $O(\log m)$ time using $\frac{nm}{\log m}$ processors.

Put together, we can compute an optimal $\boldsymbol{w}$ in

4. There are $\binom{d}{2}$ possible 2-itemsets. We can generate all possible 2-itemsets in $O(1)$ time using $\binom{d}{2}$ processors.

   For each possible 2-itemset Q we assign $\frac{n}{\log n}$ processors. The $\frac{n}{\log n}$ processors associated with Q do the following: Each processor is assigned $\log n$ transactions. The processor checks if Q is in each one of the $\log n$ transactions and counts the number of transactions in which Q is a subset. The $\frac{n}{\log n}$ processors associated with Q add the numbers generated above using a prefix sums computation. The total sum of these numbers is the number of transactions in which Q occurs. If this number is $\geq n\times$minSupport, then Q is output as a frequent 2-itemset.

   Clearly, the total number of processors needed is $\frac{n}{\log n}\binom{d}{2} \leq \frac{d^2 n}{\log n}$. The run time is $O(\log n)$.