**DSC 255: Machine learning**

# Week 10 — Solutions

1. *Neural net parametrization.* Ignoring offset terms, we have

   - $10 \times 1000 = 10{,}000$ parameters between the input layer and first hidden layer;
   - $1000 \times 1000 = 1{,}000{,}000$ parameters between successive hidden layers;
   - $1000 \times 1 = 1000$ parameters between the last hidden layer and the output layer.

   The total number of parameters is thus 3,011,000.

2. *Softmax probabilities.* The most likely label is 1, which has probability

$$\frac{e^1}{e^1 + e^0 + e^{-1} + e^0} \approx 0.53.$$

3. *Exclusive-OR.* There are many ways to do this, for instance:

   - $h_1 = \text{ReLU}(x_1 - x_2)$
   - $h_2 = \text{ReLU}(x_2 - x_1)$
   - $y = h_1 + h_2$

   which behaves as follows:

| $x_1$ | $x_2$ | $h_1$ | $h_2$ | $y$ |
|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 0 |
| 0 | 1 | 0 | 1 | 1 |
| 1 | 0 | 1 | 0 | 1 |
| 1 | 1 | 0 | 0 | 0 |

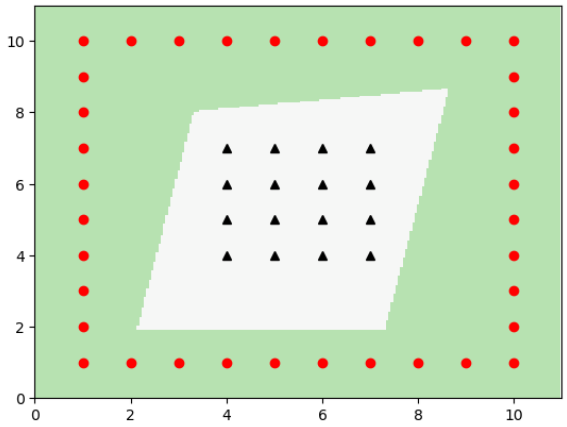
4. *Decipher the net.*

   (a) $y = x_1 \vee x_2$; that is, the neural net computes the Boolean OR function.

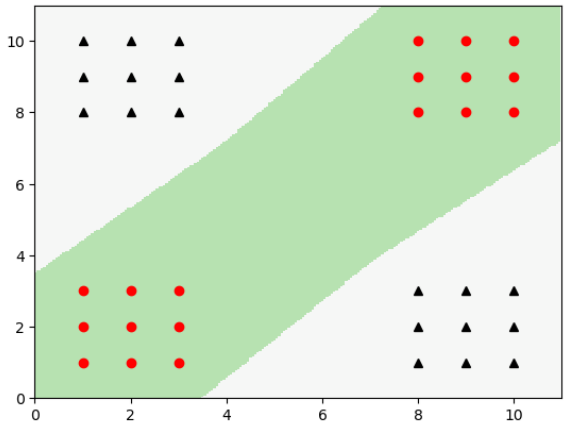
   (b) $y = |x|$; that is, the neural net computes the absolute value function.

5. *Neural net experiments.*

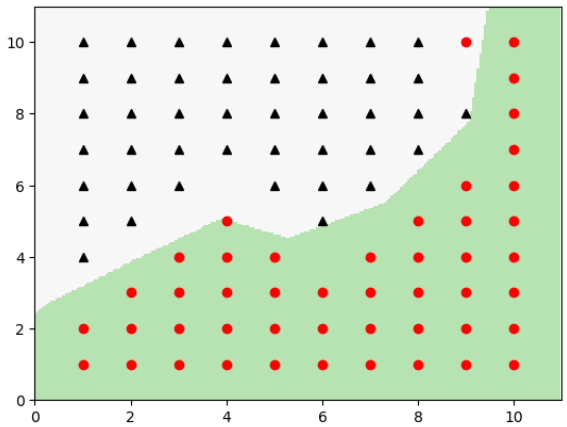
   (a) `data1.txt`: $H = 4$, number of iterations is 240K.



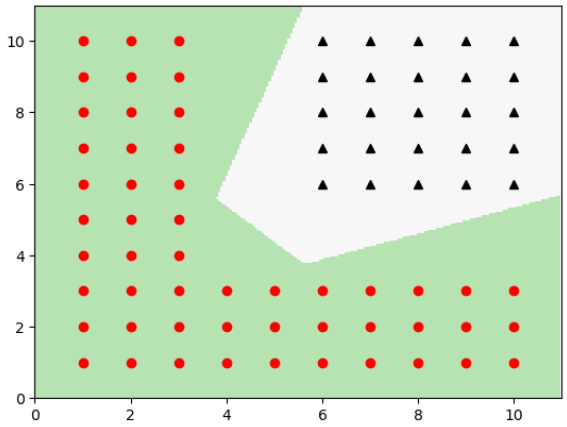

`data2.txt`: $H = 4$, number of iterations is 63K.



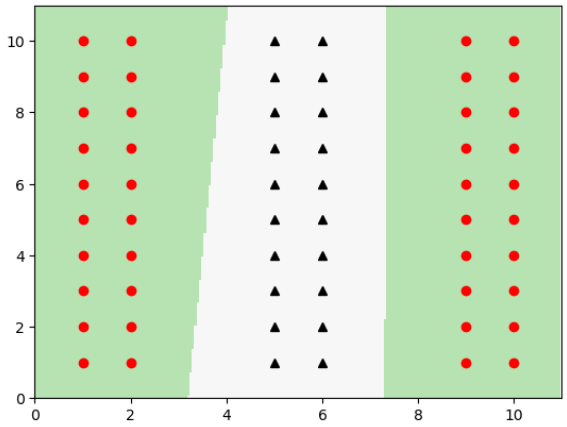

`data3.txt`: $H = 32$, number of iterations is 539K.



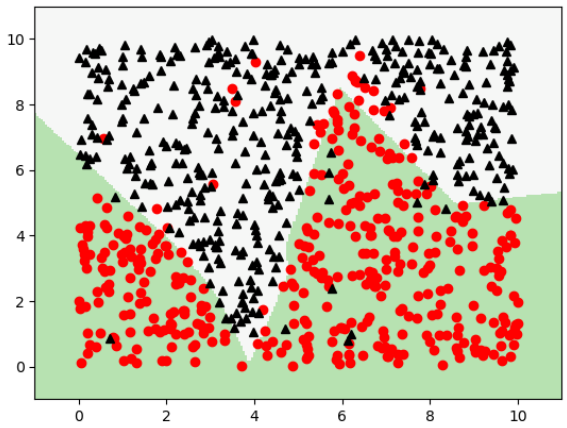

`data4.txt`: $H = 4$, number of iterations is 55K.




`data5.txt`: $H = 4$, number of iterations is 178K.

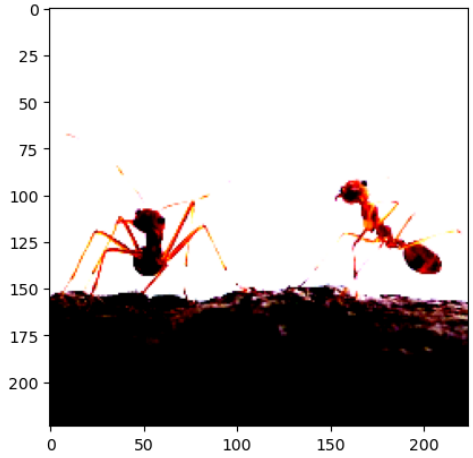



(b) This is with $H = 4$.




6. *A computer vision classification task.*

(a) Here is the sixth ant image in the training set, before and after transformation.


(b) The logistic regression classifier has test accuracy 80.4%.

(c) The $k$-nearest neighbor classifier has test accuracy $68.6\%, 71.9\%, 69.3\%$ for $k = 1, 3, 5$.