6.034 Quiz 3
15 November 2013

Name	KARN
email	karn@tolaria.edu

Circle your TA **(for 1 extra credit point)**, so that we can more easily enter your score in our records and return your quiz to you promptly.

Michael Fleder **Giuliano Giacaglia** **Dylan Holmes**

Casey McNamara **Robert McIntyre** **Duncan Townsend**

Mark Seifter **Sam Sinai** **Prashan Wanigasekara**

Problem number	Maximum	Score	Grader
1	50		
2	50		
Total	100		

There are 8 pages in this quiz, including this one, but not including blank pages and tear-off sheets. Tear-off sheets with duplicate drawings and data are located after the final page of the quiz. As always, open book, open notes, open just about everything, including a calculator, but no computers.

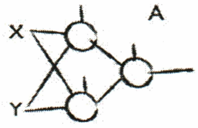
Problem 1: Neural Networks (50 points)

Part A: The Paleplane puzzle (24 points)

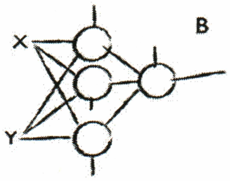
After a strange incident involving a malfunctioning quiz problem, 6.034 TA Dylan Holmes has vanished from within Prof. Winston's locked office! Detective (and distant relative) Sherlock Holmes has just broken in to look for clues.

He finds four suspicious neural nets drawn on a whiteboard:

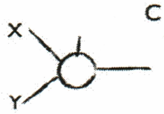
A



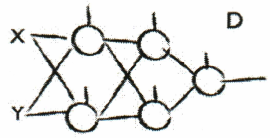
B



C

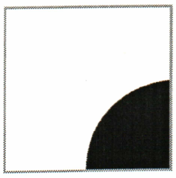


D



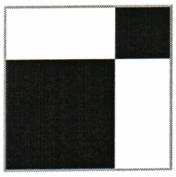
Nearby, he also finds ten ominous figures which may represent the output of some of the neural nets:

A B D



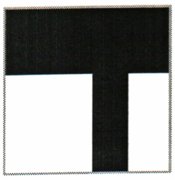
NONE

NONE

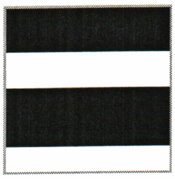


D

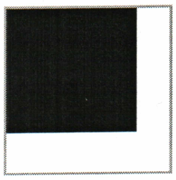
A B C D



B



B



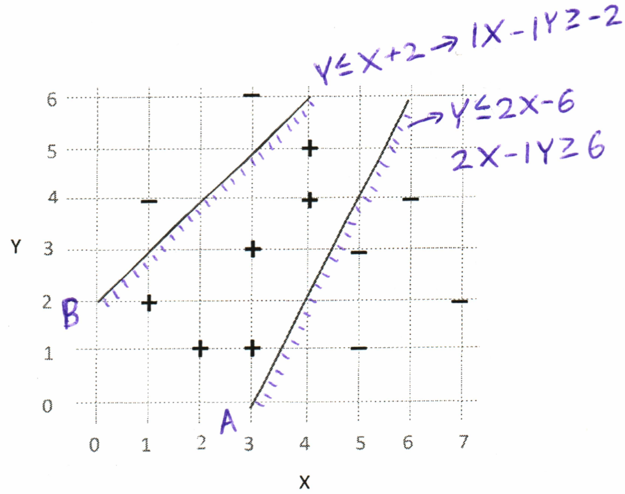
A B D

In the box below each figure, **list all** the neural nets (A, B, C, D) that could have produced that figure as output. **If none of the neural nets could produce the figure, write NONE instead**. Assume all the neural nets use the unit step threshold function, that is:

$$u(x)=\begin{cases}0 & x<T\\ 1 & x\geq T\end{cases}$$

Part B: Heads and Tails (26 points)

Suddenly, Sherlock's friend John runs into the office waving a map. It turns out that a fearsome mathematician has been using a neural net to predict Dylan's location!

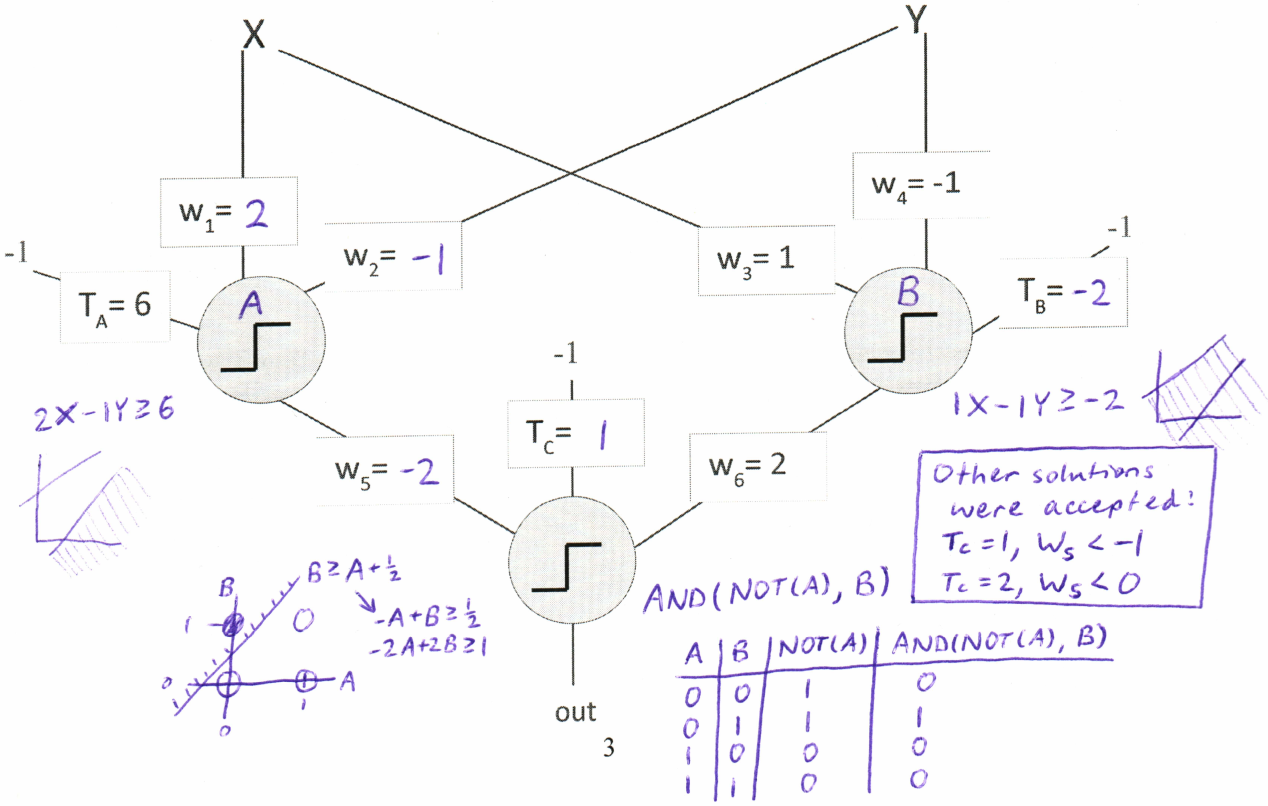


On the right is a campus map of MIT; points marked **+** are places that Dylan has visited, and points marked **–** are places that Dylan has avoided.

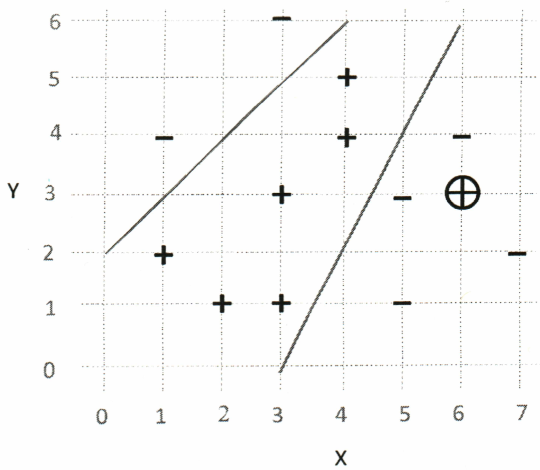
Below is the neural net that was trained to predict Dylan's location—it produces an output of 1 when its inputs lie between the two lines drawn on the map, and an output of 0 otherwise.

B1. (20 points) Find weights and thresholds for this neural net which would enable it to produce the described behavior. **Each value should be an integer**, and some of the values have been filled in for you.

Note: A copy of these diagrams, along with scratch space, is provided on a tear-off sheet after the last page of the exam.



B2. (6 points). Prof. Winston arrives, appropriately surprised to see two fictional strangers reading a map in his locked office. John explains that he just saw Dylan at the point ⟨6,3⟩ . He wants to adjust the neural net from part B1 to account for this new positive training point – but he knows that neural nets are fragile, so he doesn't want to tamper with it very much. Intrigued, Prof. Winston asks him how he would do it.



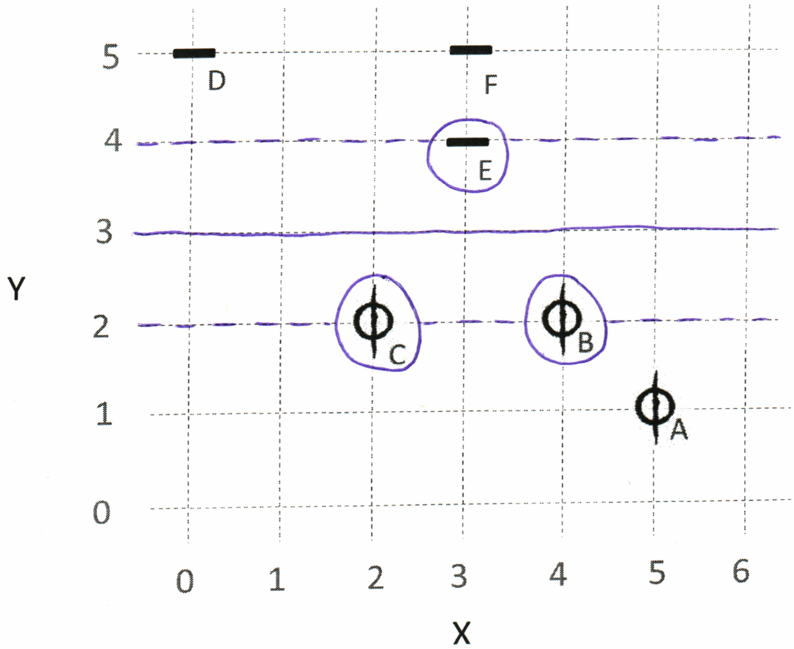
Your task is to change as few of the weights/thresholds as possible so that the neural net from part B1 correctly classifies **all** of the training points—including the new positive point at ⟨6,3⟩ . (You do **not** need to use backward propagation; you can solve by inspection.)

List a minimal set of weights and thresholds (w_1, w_2, w_3, w_4, w_5, w_6, T_A, T_B, T_C) you would need to change. If none of the weights/thresholds need to change, write "NO CHANGE". If the task is impossible, write "IMPOSSIBLE" instead.

IMPOSSIBLE. Two lines are not sufficient to separate the training data, so we need more neurons.

Problem 2: Oil vectors support machines (50 points)

The great wizard Urza has just learned that his home is being contaminated by the spread of glistening oil. He has collected several soil samples from the vicinity and marked them as either contaminated (Φ) or normal (–).



Part A (30 points):

Drawing upon his mastery of support vector machines, Urza wants to construct a linear boundary that separates the contaminated samples from the normal samples by the widest possible margin.

A1 (10 points).

On the diagram above,

1. Draw the SVM boundary as a solid line.
2. Draw the gutters as dashed lines.
3. **Circle** all of the support vectors.

A2 (20 pts). Based on the boundary you drew:

(a) Determine the width of the margin.

margin width = distance between gutters = 2

(b) Solve for the vector $\overline{w}$ and offset b. (Recall the positive side of the boundary is defined by $\vec{w}\cdot\vec{x}+b\geq 0$. Contaminated samples (Φ) should be classified as positive; normal samples (–) should be classified as negative).

$\overline{\mathbf{w}} = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$ b = 3

(c) Solve for the supportiveness values α_i.

$\alpha_A = 0$ $\alpha_D = 0$

$\alpha_B = \frac{1}{4}$ $\alpha_E = \frac{1}{2}$

$\alpha_C = \frac{1}{4}$ $\alpha_F = 0$

Show your work for partial credit:

$\vec{w}$ & b

1. Draw decision boundary line (done)
2. Write equation: $Y=3$
3. Rewrite as $\vec{w}\cdot\vec{x}+b=0$

$$[0\ \ 1]\begin{bmatrix} x \\ y \end{bmatrix} + (-3) = 0$$

4. Scale equation to remove degree of freedom using gutter constraint.

For $\vec{x}_i$ on gutter: $(\text{sign}_i)(\vec{w}\cdot\vec{x}_i + b) = 1$

$(+1)\left([0\ \ 1]\begin{bmatrix} 2 \\ 2 \end{bmatrix} - 3\right) = -1 \rightarrow$ want $+1$

Scale: multiply by -1

$$[0\ \ -1]\begin{bmatrix} x \\ y \end{bmatrix} + 3 = 0$$

$\vec{w} = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$

$b = 3$

α

1. For non-support vectors, $\alpha = 0$

$\alpha_A = \alpha_D = \alpha_F = 0$

2. $\sum \alpha_+ = \sum \alpha_-$ (positive support vectors; negative support vectors)

$\alpha_B + \alpha_C = \alpha_E$

3. $\vec{w} = \sum \alpha_+ \vec{x}_+ - \sum \alpha_- \vec{x}_-$ (positive support vectors; negative support vectors)

$$\begin{bmatrix} 0 \\ -1 \end{bmatrix} = \alpha_B \begin{bmatrix} 4 \\ 2 \end{bmatrix} + \alpha_C \begin{bmatrix} 2 \\ 2 \end{bmatrix} - \alpha_E \begin{bmatrix} 3 \\ 4 \end{bmatrix}$$

$0 = 4\alpha_B + 2\alpha_C - 3\alpha_E$

$-1 = 2\alpha_B + 2\alpha_C - 4\alpha_E$

$-1 = 2\alpha_E - 4\alpha_E \rightarrow \alpha_E = \frac{1}{2}$

$4\alpha_B + 2\alpha_C = 3(\frac{1}{2})$

$-(2\alpha_B + 2\alpha_C = 2(\frac{1}{2}))$

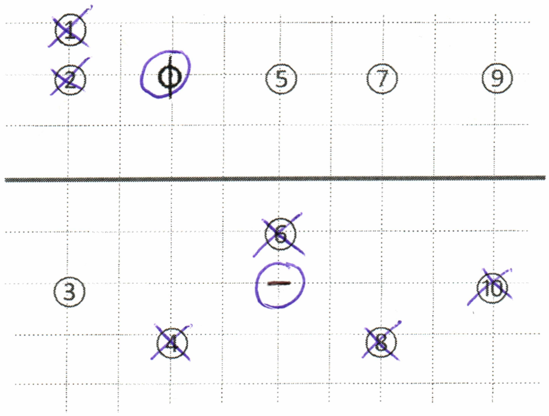
$2\alpha_B + 0 = \frac{1}{2}$

$\alpha_B = \frac{1}{4}$

$\alpha_C = \frac{1}{4}$

Part B (20 points):

B1 (10 points)



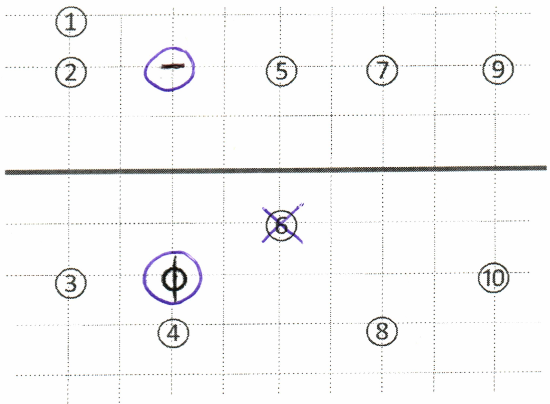
The SVM boundary in this diagram was produced by a training set consisting of just three training points. (Some of them, but not necessarily all of them, are support vectors.) Although you know the location of the two training points shown, you've forgotten the location of the third training point —and whether it was contaminated or normal!

Out of the ten locations marked ①-⑩, where could the third training point have been?

List all the locations ①-⑩ where the third training point could have been. If it could not have been in any of those locations, write NONE instead.

3, 5, 7, 9

B2 (10 points). Suppose the first two training points were arranged like this instead: Out of the ten locations marked ①-⑩, where could the third training point have been?



List all the locations ①-⑩ where the third training point could have been. If it could not have been in any of those locations, write NONE instead.

1, 2, 3, 4, 5, 7, 8, 9, 10

Problem 3: Spiritual and Right-Now (5 points)

Circle the **best** answer for each of the following questions. There is **no penalty for wrong answers**, so it pays to guess in the absence of knowledge.

1 Boyden indicated it is now possible to:

1. Record neuron firings using ion-sensitive dyes.
2. Record neuron firings in cell DNA.
3. Record neuron firings in microcircuits embedded in cells.
4. Record neuron firings by embedding light-emitting diodes in cell walls.
5. Record neuron firings by connecting sensors to neurons with artificial synapses.

2 Katz's START system was instrumental in the development of:

1. YouTube.
2. Facebook and twitter.
3. Mechanical turk.
4. Siri and Watson.
5. Big data research.

3 Genetic algorithms involve:

1. Distance computation using vector dot product.
2. A mechanism that helps to avoid problems with local maxima.
3. Few programmer choices, hence easy implementation.
4. Felicity conditions.
5. Mapping from learned skills to analogs of animal geneotype.

4 Learning phonological rules using the Sussman-Yip method:

1. Uses positive examples only.
2. Exploits the McGurk effect.
3. Starts from a "seed" example which it generalizes.
4. Starts from a "seed" example which it specializes.
5. Uses nearest-neighbors distances to make guesses.

5 Transition space:

1. Marks the period when teenagers start to think about life away from home.
2. Suggests that in human thinking, change causes change.
3. Reflects the idea that a system's state determines its future history.
4. Features a vocabulary of roles, such as agent, instrument, and object.
5. Was conceived as a way to account for human linguistic competence.

6 Near miss learning involves

1. Learning, with the benefit of big data algorithms.
2. Learning, with the number of samples approximating the number of characteristics learned.
3. Learning, with statistical regularity governing what is learned.
4. Learning, without supervision.
5. Learning, without a symbolic representation.