6.034 Quiz 3

9 November 2016

Name	Leif Erikson
Email	

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

Jake Barnwell	**Michaela Ennis**	**Rebecca Kekelishvili**
Vinny Chakradhar	**Phil Ferguson**	**Nathan Landman**
Alex Charidis	**Stevie Fine**	**Samarth Mohan**
Brian Copeland	**Brittney Johnson**	**Jessica Noss**

Problem	Maximum	Score	Grader
1 - Neural Nets	50		
2 - SVM	50		
Total	100		

SRN	6		

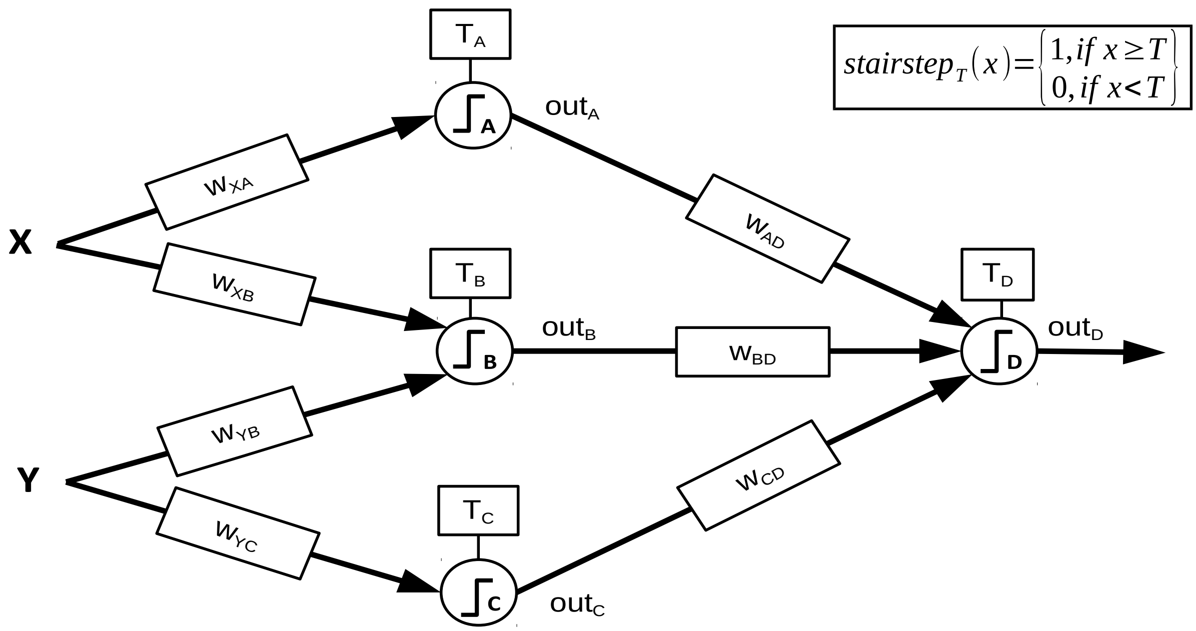
There are 13 pages in this quiz, including this one.

As always, open book, open notes, open just about everything, including a calculator, but no computers or phones.

Problem 1: Neural Networks (50 points)

Part A: Forward Propagation (40 points)

Disni Corporation is coming out with a new animated motion picture. The working title is: *How to train your network*. They ask you to help them work out the technical details of their script, so as to not embarrass themselves in front of Gaffry Hilton, their most devoted fan. They start with the neural net below:



A1 (9 points) For each of the neurons A, B, and C, write the equation for the neuron's output in terms of the inputs (X, Y), weights (e.g. W_{XA}), and thresholds (e.g. T_A) in the neural net.

$$out_A = \begin{cases} 1, \text{ if } \boxed{W_{XA} \cdot X \geq T_A} \\ 0, \text{ otherwise} \end{cases}$$

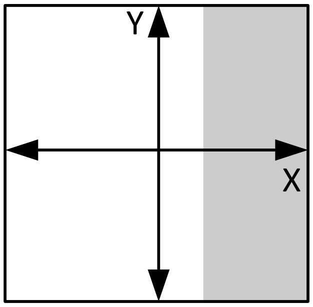
$$out_B = \begin{cases} 1, \text{ if } \boxed{W_{XB} \cdot X \cdot W_{YB} \cdot Y \geq T_B} \\ 0, \text{ otherwise} \end{cases}$$

$$out_C = \begin{cases} 1, \text{ if } \boxed{W_{YC} \cdot Y \geq T_C} \\ 0, \text{ otherwise} \end{cases}$$

A2 (6 points) Consider the **output of Neuron B (out$_B$)**. By changing the thresholds and weights in the neural net from the previous page, which of the following pictures could Neuron B be trained to draw? Next to each picture, circle **YES** if Neuron B could draw the picture, otherwise circle **NO**.

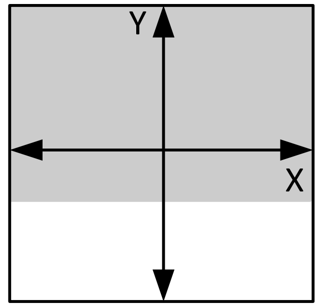
YES

NO



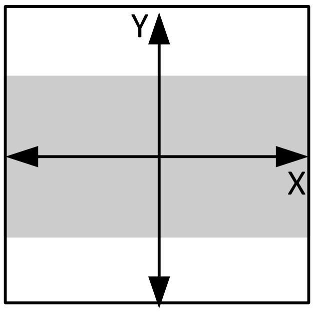
YES

NO



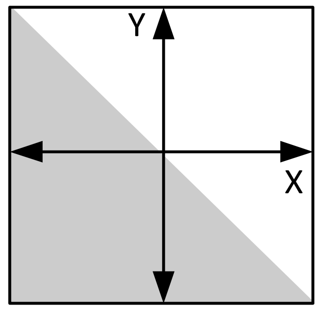
YES

NO



YES

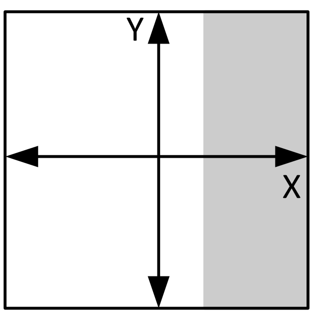
NO



A3 (6 points) Now consider **Neuron C** instead and answer the same question.

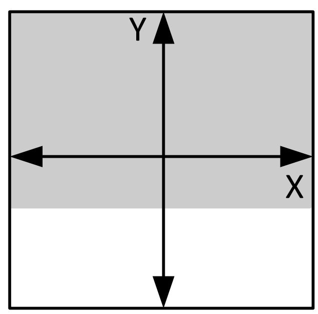
YES

NO



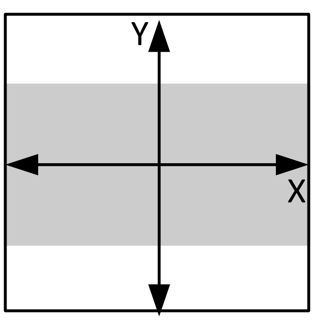
YES

NO



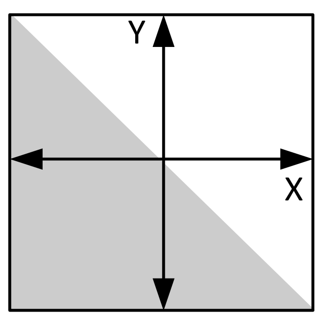
YES

NO

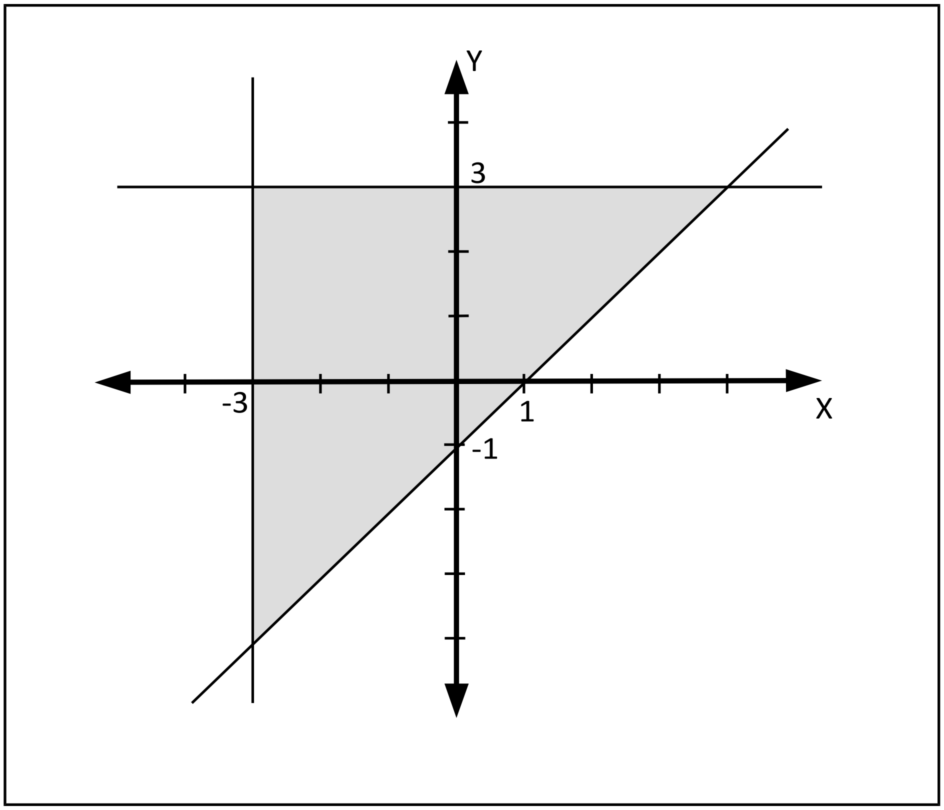


YES

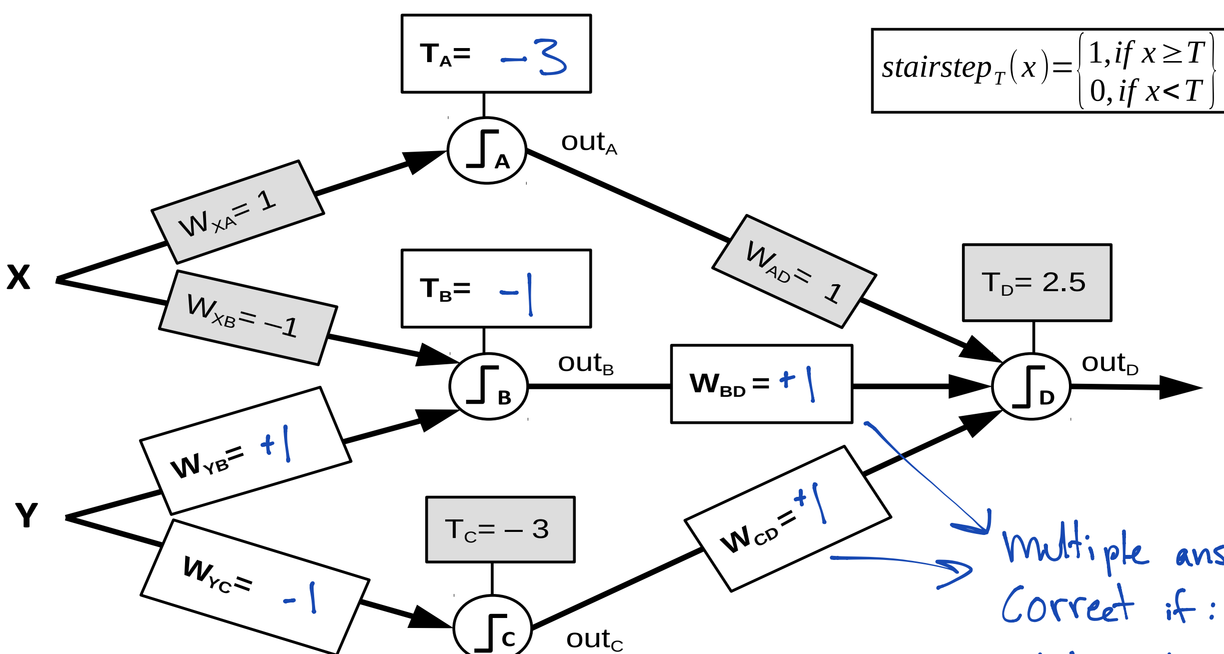
NO



A4 (19 points) Disni assigns some weights and thresholds to the neural network. As shown in the plot below, the network produces an output of 1 in the shaded (gray) area, and an output of 0 in the unshaded (white) area.



In the neural net below, fill in the weights and thresholds. Some of the weights and thresholds have been provided for you. Fill in $\mathbf{w_{YB}}$, $\mathbf{w_{YC}}$, $\mathbf{T_A}$, $\mathbf{T_B}$, $\mathbf{w_{BD}}$, and $\mathbf{w_{CD}}$. (For partial credit, show your work **on the next page**.)



multiple answers;
Correct if:

$W_{BD} + W_{CD} \geq 1.5$

$W_{BD} < 1.5$

$W_{CD} < 1.5$

Show your work for partial credit:

Use line equations for first layer

$x \geq -3 \rightarrow 1x + 0y \geq -3 \rightarrow W_{xA} = 1, T_A = -3$

$y \leq 3 \rightarrow 0x + -1y \geq -3 \rightarrow W_{yC} = -1, T_C = -3$

$y \geq 1x - 1 \rightarrow -1x + 1y \geq -1 \rightarrow W_{xB} = -1, W_{yB} = +1, T_B = -1$

AND of neurons A, B, & C for last layer

if all true:

$1 \cdot W_{AD} + 1 \cdot W_{BD} + 1 \cdot W_{CD} \geq 2.5 \rightarrow \boxed{W_{BD} + W_{CD} \geq 1.5}$

if one not true:

$W_{AD} + W_{BD} < 2.5 \rightarrow \boxed{W_{BD} < 1.5}$

$W_{BD} + W_{CD} < 2.5$

$W_{AD} + W_{CD} < 2.5 \rightarrow \boxed{W_{CD} < 1.5}$

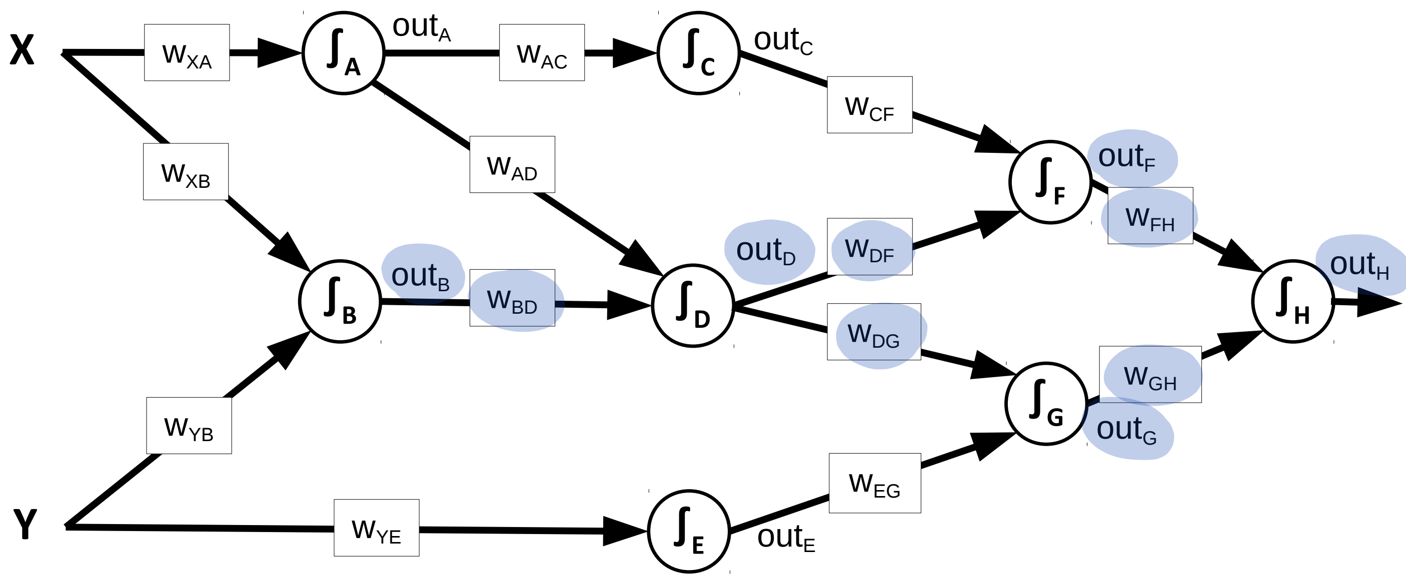
Choose $W_{BD} = W_{CD} = 1$

Part B: Backward Propagation (10 points)

To create a movie scene featuring neural network training, Disni producers introduce the new neural network pictured below, in which neurons use the sigmoid function instead of the stairstep function.

The producers have computed the output of each neuron using the sigmoid function (for example, out_A is the computed output of neuron A). Now, they want to know what information is needed to **update the weight w_{BD} going from neuron B to neuron D**. Given that they have already calculated the outputs, which inputs (e.g. Y), weights (e.g. w_{XA}), and neuron outputs (e.g. out_A) will the producers need in order to compute the **new weight w_{BD}**?

On the neural net below, **circle each input, weight, and neuron output that is required**.



Show your work for partial credit:

$$W_{BD}' = W_{BD} + r \cdot out_B \, \delta_D$$

$$\delta_D = out_D (1 - out_D) \cdot (W_{DF} \, \delta_F + W_{DG} \, \delta_G)$$

$$\delta_F = out_F (1 - out_F)(W_{FH} \, \delta_H) \qquad \delta_G = out_G (1 - out_G)(W_{GH} \, \delta_H)$$

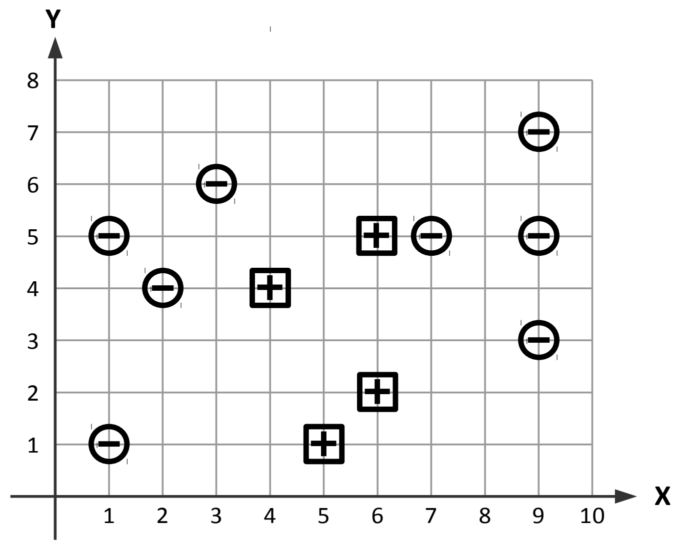
$$\delta_H = out_H (1 - out_H)(out^* - out_H)$$

Problem 2: Support Vector Machines (50 points)

Thanksgiving is coming! You want to host a Thanksgiving party at your apartment. In the spirit of Thanksgiving, you want to cook a **TURKEY** for the meal. However, when you get to Costco, you find that the turkeys are mixed in with the **CHICKEN**s.

Part A: Picking a Model (6 points)

You want to create a model to correctly classify turkeys and chickens. Because you are an experienced zoologist, you can easily tell the turkeys apart from the chickens. Below, they are shown on a grid with axes **Y** and **X**, where **Y** represents the shelf height and **X** represents position along a shelf. In this grid, **turkeys (⊞) are positive samples** and **chickens (⊖) are negative samples.**



Which of the following machine-learning models could be used to classify this data? (That is, for each model given, is it possible to train a model of that form to classify the data perfectly?)

Of the six (6) models given below, circle ALL that apply. If none apply, instead circle **NONE OF THESE**.

Neural Net with exactly 2 neurons

SVM with a Linear Kernel

Neural Net with exactly 3 neurons

SVM with a Quadratic Kernel

Neural Net with exactly 4 neurons

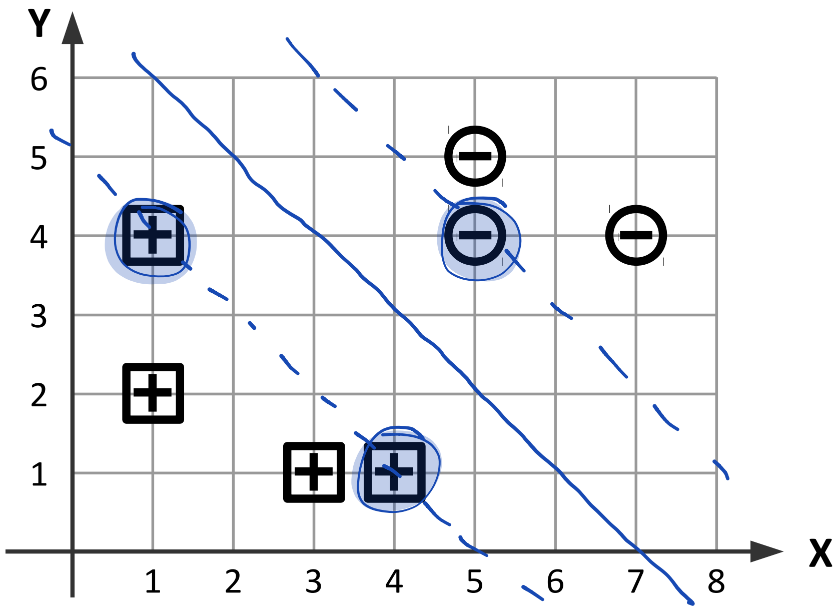
SVM with a Radial Basis Function Kernel

NONE OF THESE

Part B: Cold Turkey (18 points)

While you were solving the previous problem, a customer purchased many of the turkeys and chickens, then rearranged the remaining ones! Fortunately, you realize the new arrangement of turkeys and chickens can be classified using an **SVM with a linear kernel**.

B1 (10 points) The new arrangement of turkeys and chickens is shown in the graph below. As before, **turkeys (⊞) are positive samples** and **chickens (⊖) are negative samples**.



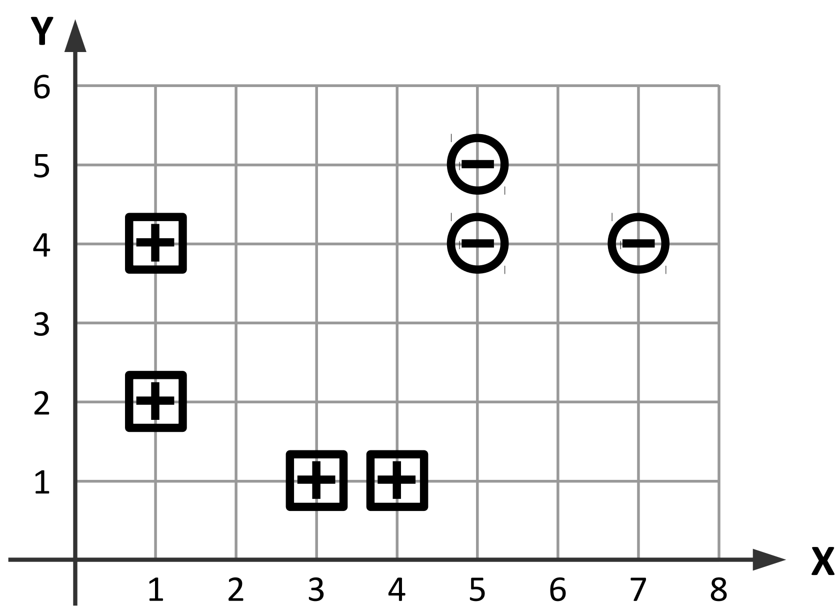
On the graph above,

- draw the linear SVM boundary with a **solid line**,
- draw the positive and negative gutters with **dashed lines**,
- and **circle** all support vectors.

If you want to redraw your solution, use this graph and CHECK THE BOX BELOW.

(If you don't check the box or clearly indicate which copy you want graded, we will grade the copy above.)

☐ **Grade this copy**

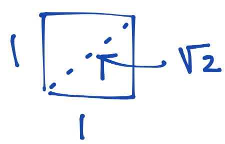


B2 (4 points) Using the boundary you drew, you can easily compute $\|\vec{w}\|$ (that is, the length of the vector $\vec{w}$) using a *geometric* solution. What is $\|\vec{w}\|$? **Fill in the box below** with a number or numeric expression.

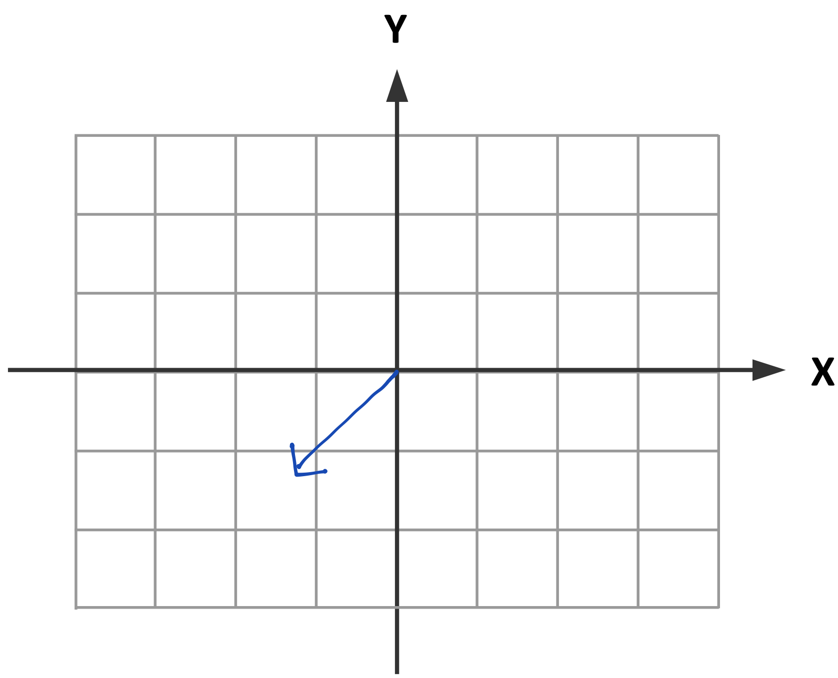
$$\|\vec{w}\| = \boxed{\frac{1}{\sqrt{2}}}$$

Show your work for partial credit:

$$\|\vec{w}\| = \frac{2}{\text{margin width}} = \frac{2}{2\sqrt{2}} = \frac{1}{\sqrt{2}} = \frac{\sqrt{2}}{2}$$

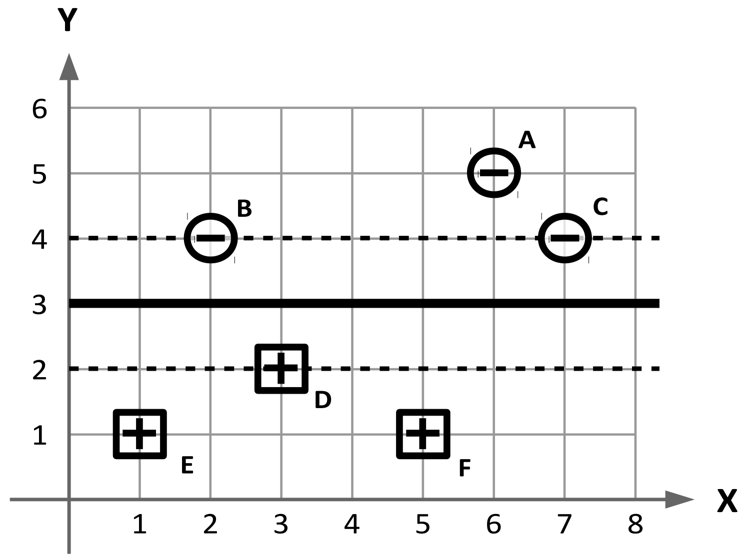


B3 (4 points) Based on the boundary you drew in Part B1, **in what direction does the vector $\vec{w}$ point?** On the graph below, **draw a vector** from the origin pointing in the correct direction, with any reasonable length. (For this problem, you do **not** need to calculate the actual value or length of $\vec{w}$.)



Part C: Supporting other Supermarkets (16 points)

While you were finishing the previous problem, a less analytical customer came by and purchased all of the remaining poultry, so you decide to head to Sam's Club instead. At Sam's Club, you find that **the turkeys and chickens are arranged differently**, as shown below. We have drawn the SVM boundary and gutters for you, and have labeled all of the data points with letters for future reference. (Note that we have not circled the support vectors.)



C1 (6 points) Consider the α ("supportiveness") values for each of the points (A, B, C, D, E, F). In the table below, fill in which α values are **zero (α = 0)**, which ones are **positive (α > 0)**, which ones are **negative (α < 0)**, and **how you know**. If there are no such points, write **NONE** in the box, and **explain why there are none**. (For this problem, you do *not* need to calculate the exact α values.)

Each of the points (**A**, **B**, **C**, **D**, **E**, **F**) should appear exactly once.

List all points for which...		Justification (10 words or fewer for each):
...α = 0:	A E F *(or write **NONE** if no such points)*	They aren't SVs
...α > 0:	B C D *(or write **NONE** if no such points)*	They are SVs
...α < 0:	None *(or write **NONE** if no such points)*	α always ≥ 0

C2 (4 points) Of the relations written below, **circle ALL that are true**. If none of the statements are true, instead circle **NONE OF THESE**:

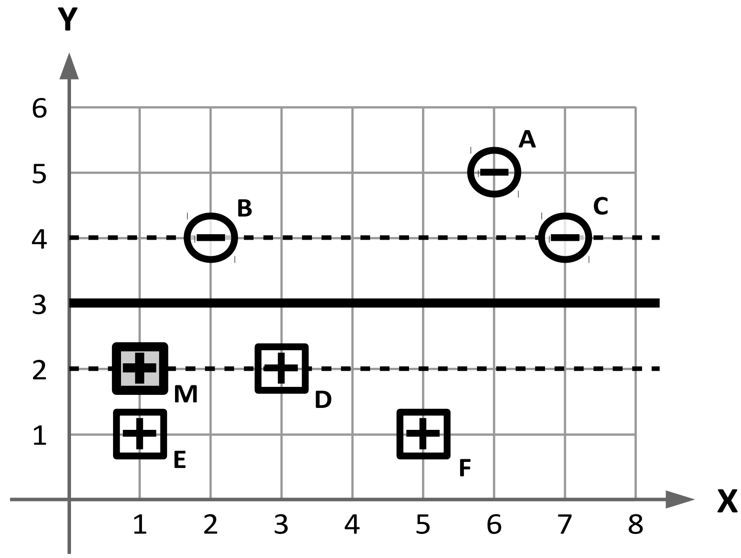
$\alpha_C < \alpha_D$ $\qquad$ $\alpha_B < \alpha_C$ $\qquad$ $\alpha_C < \alpha_B$ $\qquad$ $\alpha_D < \alpha_A$ $\qquad$ **NONE OF THESE**

C3 (6 points) For each of the three unknown points listed below, **circle the ONE best answer indicating how the SVM would classify that point**:

(x = 4, y = 5):	**TURKEY (⊞)**	**CHICKEN (⊖)**	**CAN'T TELL**
(x = 4, y = 4):	**TURKEY (⊞)**	**CHICKEN (⊖)**	**CAN'T TELL**
(x = 4, y = 3.5):	**TURKEY (⊞)**	**CHICKEN (⊖)**	**CAN'T TELL**

Part D: Changing Support (10 points)

On a whim, you decide to **add a new TURKEY (⊞) training point M at (x=1, y=2)** to the previous training data set. For your convenience, shown below is the previous SVM with the new training point (shaded and bolded) added:



Which of the following quantities ***may*** change, or remain the same, after retraining the SVM with the new training point?

For **each entry** in the table below,

- write **YES** if the value *could possibly* behave in the way specified after retraining the SVM, or
- write **NO** if the value *certainly could not* behave in the way specified.

Fill in *every* entry with either **YES** or **NO** (or **Y** or **N**, as long as you **write clearly**).
Any entry that is not clearly labeled as YES or NO (or Y or N) will be marked as incorrect.

	Could **Increase (↑)**	***Could*** **Decrease (↓)**	***Could*** **remain unchanged**
α_A	No	No	Yes
α_B	Yes	No	Yes
α_C	No	Yes	Yes
α_D	No	Yes	Yes
$\alpha_B + \alpha_C$	No	No	Yes

Problem 3: Spiritual and Right Now

Circle the **one 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. Spelke's experiment involving a rectangular room and a blue wall showed that:
    1. Humans can be reduced to rat-level performance when they repeat words they hear.
    2. Human children are good at using two sources of information at once at an early age.
    3. Rats become disoriented when sounds come from multiple directions.
    4. Human children are more likely to search for toys near brightly colored walls.
    5. Both rats and humans are confused by rapidly changing wall colors.

2. The Yip-Sussman approach to learning phonological rules:
    1. Specializes and generalizes candidate patterns using 1-nearest neighbors on end phonemes.
    2. Cannot distinguish between neighboring sounds such as "ba" and "ga".
    3. Uses deep neural nets to place decision boundaries in phoneme space.
    4. Employs a small lexicon of irregularly spelled words.
    5. Stops generalizing when an evolving pattern matches a negative example.

3. Near-miss learning:
    1. Interpolates between examples to infer missing training points.
    2. Can only learn from positive examples of the concept being taught.
    3. Starts with large groups of positive and negative examples, then combines similar groups.
    4. Starts with ordinary neural net training, which is then refined by fresh examples.
    5. Starts with a seed example and then makes modifications to both limit and expand matches.

4. A case-frame representation, also known as a role-frame representation:
    1. Associates problem cases with a small number of methods to solve them.
    2. Has slots for entities participating in events.
    3. Has been filled reliably by deep neural nets looking at images of actions such as give and drop.
    4. Has slots for values that describe state changes as inputs, and the next state as an output.
    5. Is universal because it can be used as an assembly language for all other representations.

5. Search using the genetic-algorithm idea works best if:
    1. Mutation rate is inversely proportional to crossover rate.
    2. Survival fitness increases monotonically with phenotype complexity.
    3. Diversity is a component of fitness determination.
    4. The mapping from genotype to phenotype is one-to-one.
    5. Fitness is measured on a logarithmic scale.

6. Katz is working to improve natural-language understanding using:
    1. Computer vision.
    2. Constraint propagation.
    3. Deep neural nets.
    4. Bi-gram Bayesian statistics.
    5. Near-miss learning.