6.034 Quiz 3

6 November 2019

Name	
Email	

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

Sydney Gibson	**Rui Li**	**Allison Tam**
Udgam Goyal	**Jennifer Madiedo**	**Héctor Vazquez**
Jenna Hong	**Jack Murphy**	**Eric Wong**
Damon Jones	**Mira Partha**	

Problem number	Maximum	Score	Grader
1 – Neural Networks	50		
2 - SVMs	50		
Total	**100**		

SRN	6		

There are **12 pages** in this quiz, including this one, but not including tear-off sheets. Tear-off sheets with duplicate plots and networks are located after the final page of the quiz. **We do not collect tear-off sheets, so please show your work on the quiz pages, not the tear-off sheets.**

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

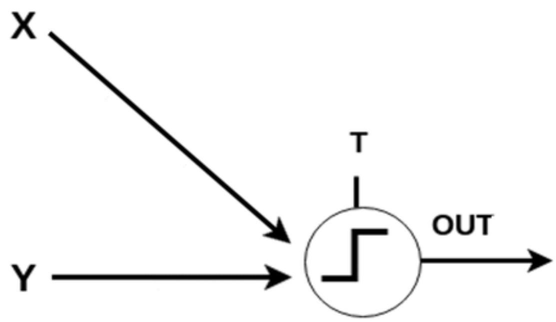
Problem 1: Neural Networks (50 points)

The 6.034 staff decides to create a neural network to help them decide whether a problem makes it into the final this year. They have access to all previous years' final exams and all the unused problems.

The neural network takes as input the characteristics of a sample problem, for example, X = length of problem, Y = difficulty of problem. **The network outputs a 1 if the problem should be included and a 0 if the problem should not be included in this year's final.**

Part A: Forward Propagation (14 points)

A1 (8 points) Consider the simple neural net below with the **stairstep activation function**:

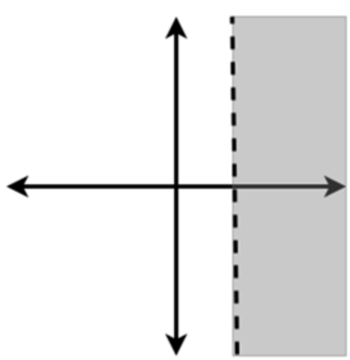


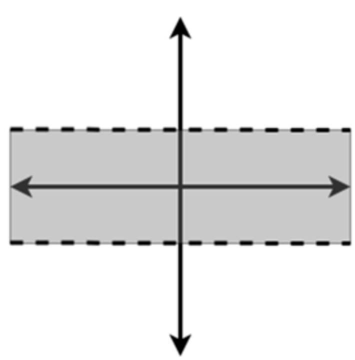
What decision boundaries can this network create? For each of the following graphs, **circle the one correct answer.**

YES

NO

CAN'T TELL





YES

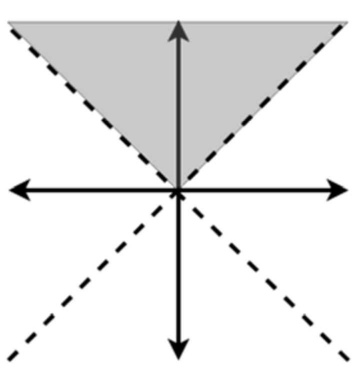
NO

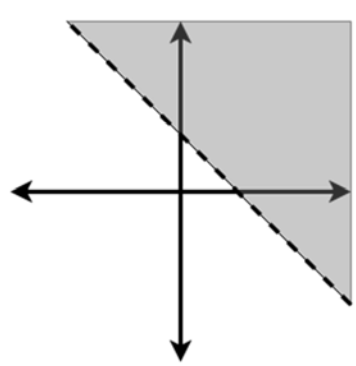
CAN'T TELL

YES

NO

CAN'T TELL





YES

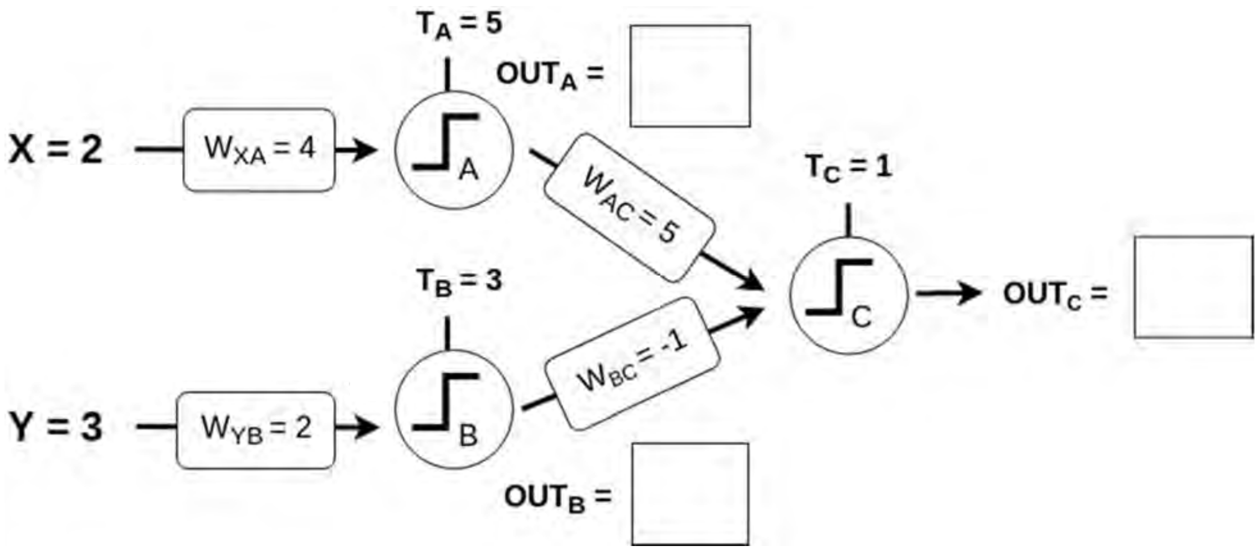
NO

CAN'T TELL

A2 (6 points) The staff creates the network shown below. Given the input **(X=2, Y=3)** to this new network, determine what the output of each neuron (**OUT_A, OUT_B, and OUT_C**) will be. **Fill in the corresponding blanks in the diagram**.

Note: Every neuron uses the **stairstep** activation function.

For your convenience, a copy of the network is provided on a tear-off sheet.



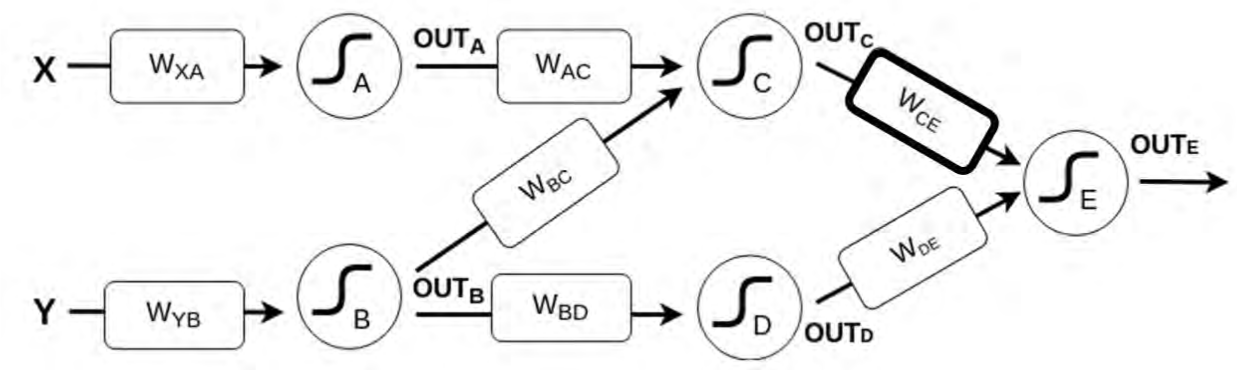
For partial credit, show your work here.

Part B: Backward Propagation (27 points)

Tired of manually fixing the weights of the network, the 6.034 staff decides to use backward propagation. The staff changes the neural network to the network shown below. In this new network, the staff has chosen the **sigmoid activation function for all the neurons**.

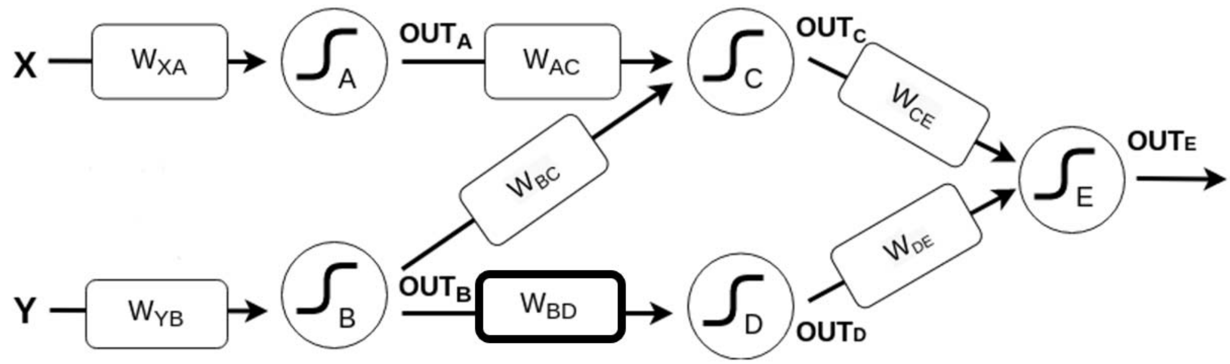
For your convenience, a copy of the network is provided on a tear-off sheet.

B1 (7 points) Suppose we want to update W_{CE} during backpropagation. On the diagram below, **circle all neuron outputs ($OUT_{A\text{-}E}$), inputs (X, Y), and weights** whose values will be used.



B2 (8 points) In the box below, write the update equation for W_{CE}. Final answers should be in terms of the neuron outputs ($OUT_{A\text{-}E}$), weights, the target output out^*, and the learning rate r.

B3 (7 points) Suppose we want to update W_{BD} during backpropagation. On the diagram below, **circle all neuron outputs (OUT_{A-E}), inputs (X, Y), and weights** whose values will be used.



B4 (5 points) Which of the following choices is the expression for calculation of ΔW_{BD}?
Circle the one best answer.

A. $\Delta W_{BD} = \frac{1}{2}(out^* - out_D)^2$

B. $\Delta W_{BD} = r * out_B[out_D(1 - out_D)W_{DE}]$

C. $\Delta W_{BD} = r * out_B[out_D(1 - out_D)(W_{DE} * [out_E(1 - out_E)(out^* - out_E)])]$

D. $\Delta W_{BD} = r * out_B[out_D(1 - out_D)(W_{DE} + W_{CE}) * [out_E(1 - out_E)(out^* - out_E)])]$

E. $\Delta W_{BD} = r * out_B\big[out_B(1 - out_B)\big(W_{XB} * X(1 - X) + W_{YB} * Y(1 - Y)\big)\big]$

For partial credit, show your work here.

Part C: True/False (9 points)

For each of the following statements, circle **TRUE** or **FALSE**.

The learning rate:

1. Is used during forward propagation.	**TRUE**	**FALSE**
2. Is used during backward propagation.	**TRUE**	**FALSE**
3. Controls the magnitude of the changes in weights.	**TRUE**	**FALSE**

Training a network:

4. Involves an iterative process of forward and backward propagation.	**TRUE**	**FALSE**
5. Could be used for supervised learning tasks.	**TRUE**	**FALSE**
6. Guarantees finding the optimal setting of weights after running for a long time.	**TRUE**	**FALSE**

Suppose that we have a neural network that is overfitting the training dataset. To help prevent overfitting, we can:

7. Stop training the neural network earlier.	**TRUE**	**FALSE**
8. Remove some layers of the network.	**TRUE**	**FALSE**
9. Remove a few neurons in each layer of the network during training.	**TRUE**	**FALSE**

Problem 2: Support Vector Machines (50 Points)

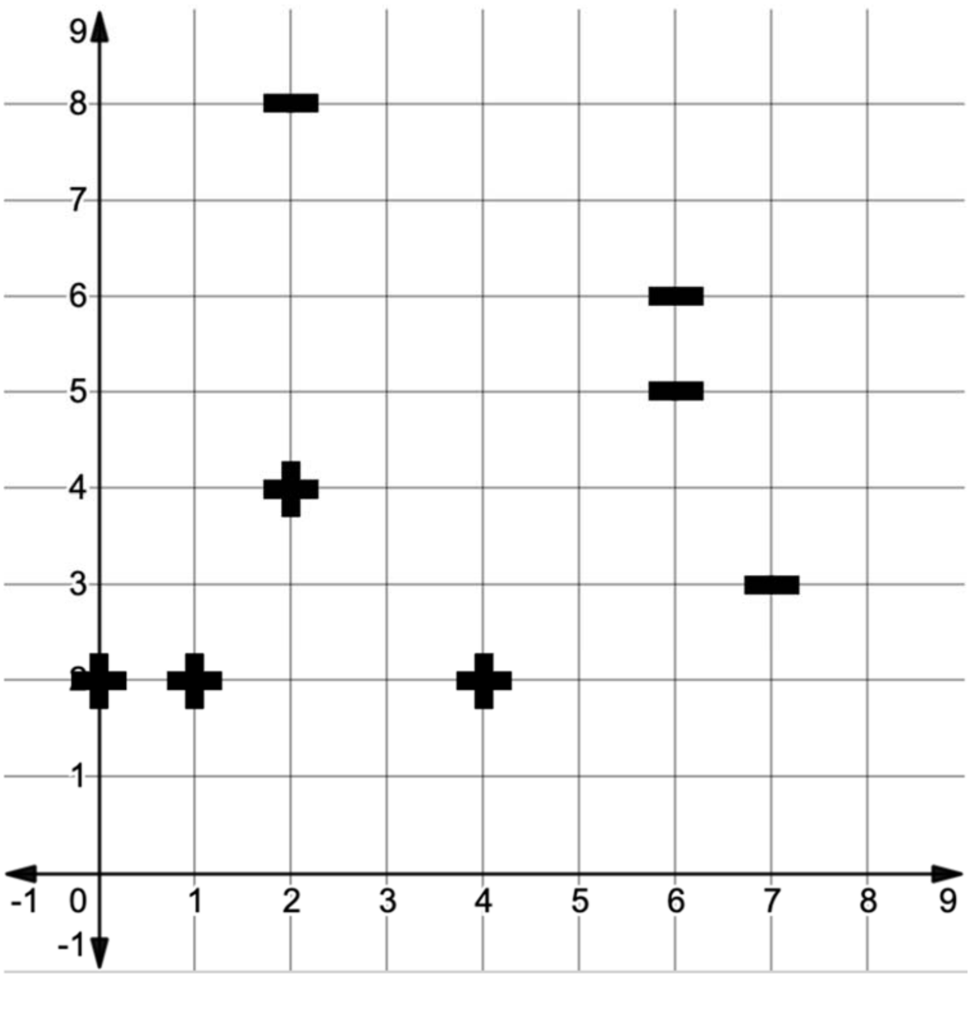
Part A (26 points): Party Planning

Ben Bitdiddle is excited to host a party this year for his **college (+)** friends and **high school (-)** friends. While deciding on dinner music, he realizes that his high school friends want to hear Christmas music and his college friends don't. Help Ben Bitdiddle figure out how to partition the dinner area into two different music zones.

A1 (8 points) Ben Bitdiddle decides to use a **Linear SVM** to partition the dinner area to separate his **college (+)** friends from his **high school (-)** friends. The diagram below shows the current locations for his friends. On the plot:

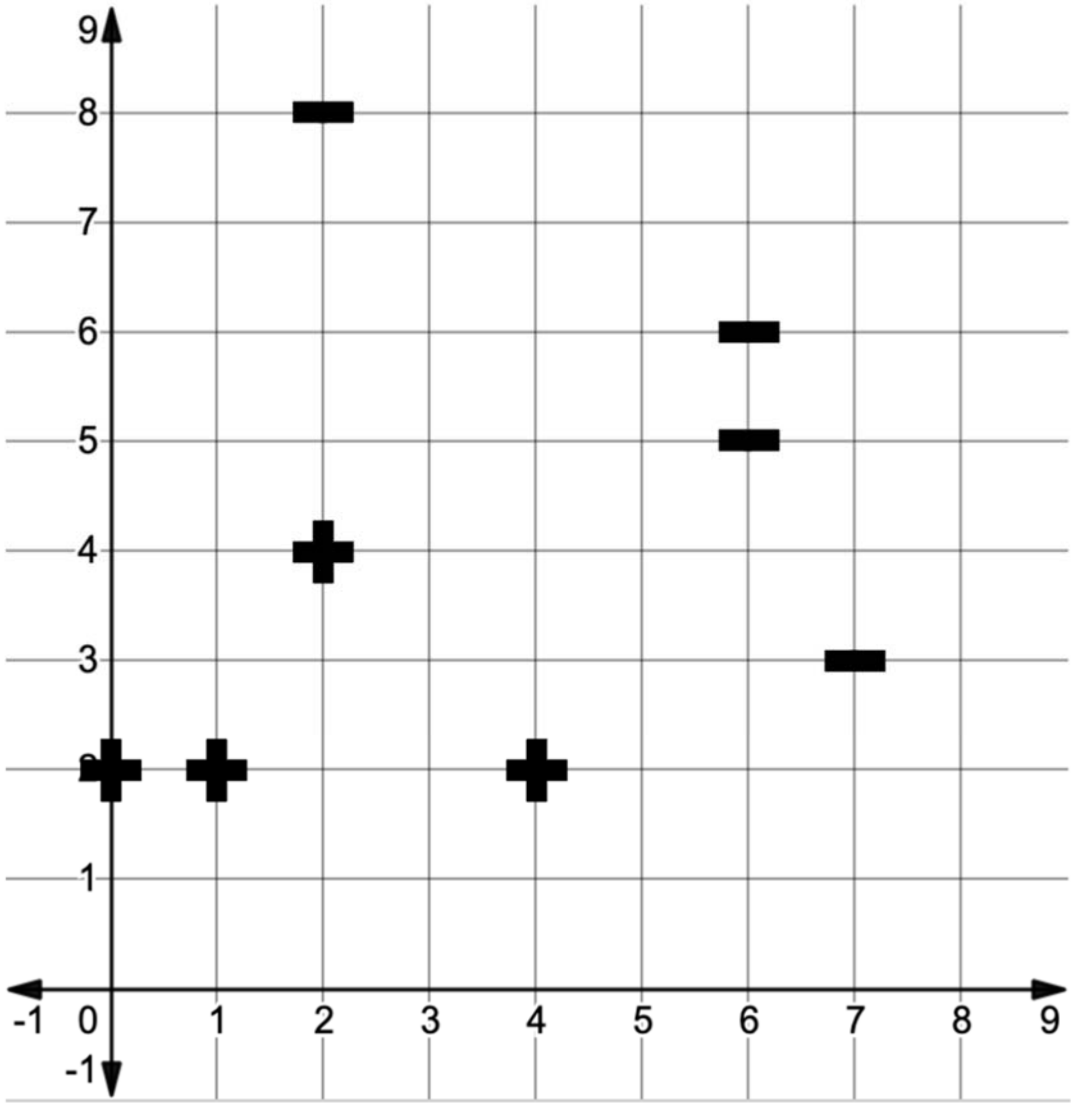
- Draw the SVM **decision boundary** with a solid line.
- Draw the SVM **gutters** with dotted lines.
- Circle the **support vectors**.

(If you want to start over, use the plot on the next page. If you write on both pages, clearly mark which page we are to grade.)



This is a duplicate copy of the plot for A1. If you want this copy graded instead, check the box.

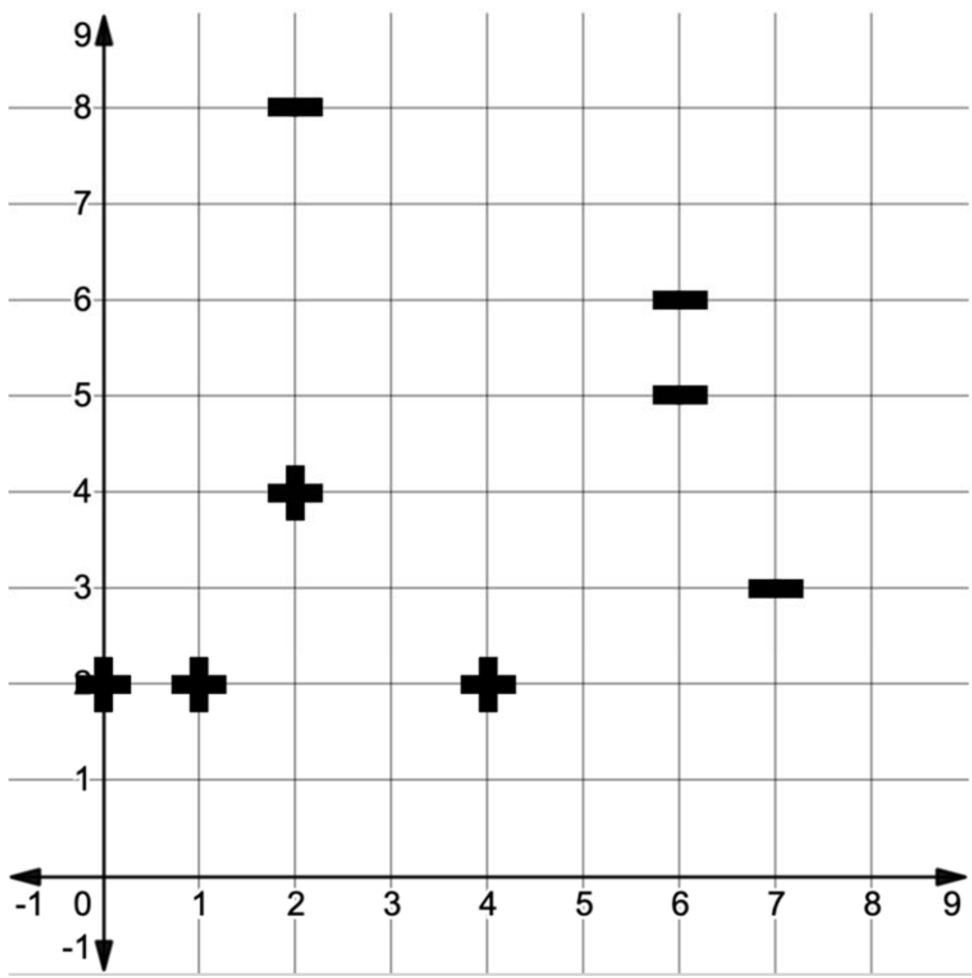
I want to start over; grade this copy



A2 (2 points) Compute the width of the margin on the decision boundary you drew in part A1. You can leave this value unsimplified, containing decimals and square roots.

A3 (16 points) Ben Bitdiddle wants to be prepared in case someone doesn't show up. **Circle each point whose removal would change the decision boundary or margin .**

The plot from A1 is reproduced below for your reference.

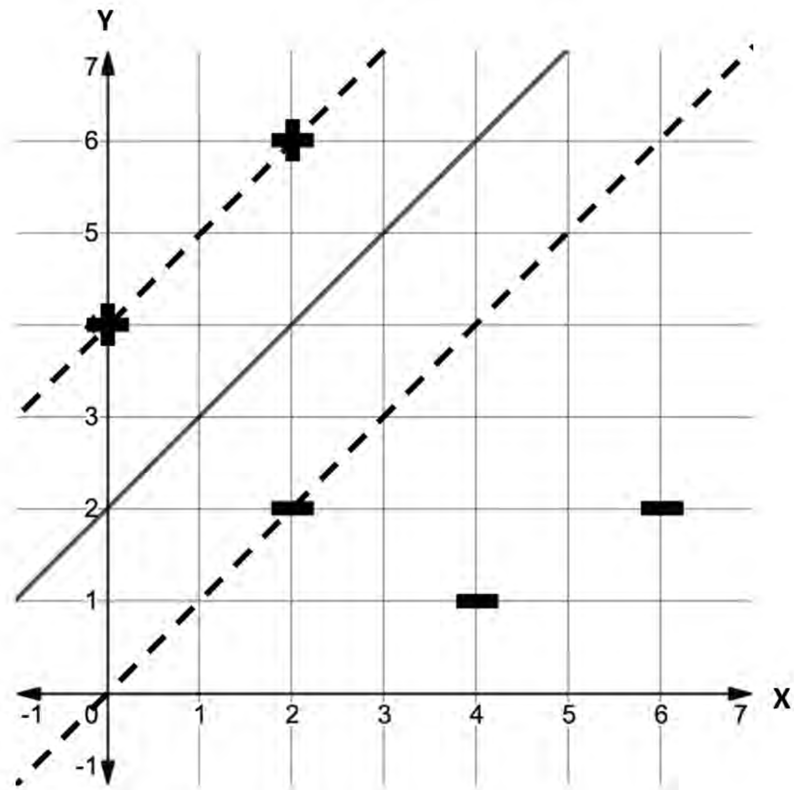


Part B (16 points): RSVP Changes

Ben Bitdiddle is finalizing the guest list after people let him know if they will attend or not. He formulates a new seating plan, shown below, again to partition the space into two music zones based on separating his **college (+)** from his **high school (-)** friends**.**

B1 (10 points) Using the given decision boundary and data points shown below, compute the values of the normal vector $\vec{w}$ and offset *b*, and write the values in the boxes below.

For your convenience, a copy of the plot is provided on a tear-off sheet.



$\vec{w}$ = ☐ b = ☐

For partial credit, show your work here.

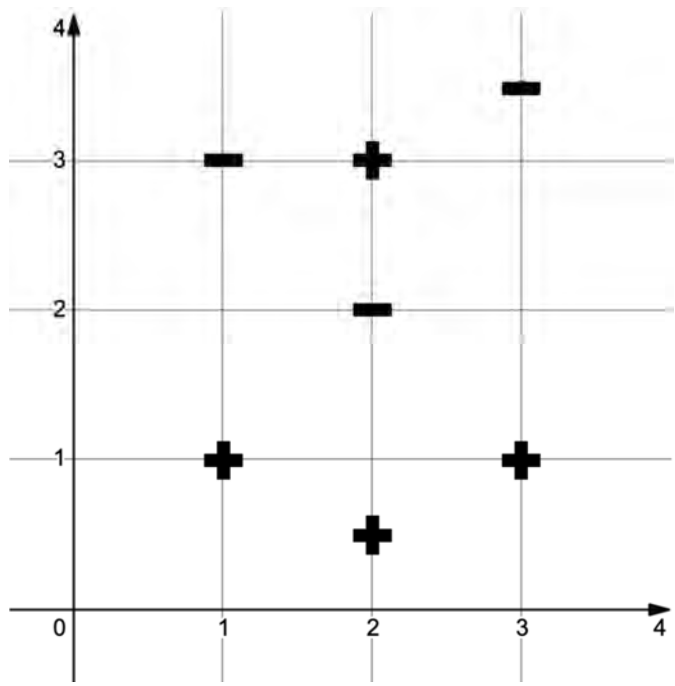
B2 (6 points) How would the SVM classify the following points?

(2, 4)	College (+)	High School (-)	Can't Tell
(0, 10)	College (+)	High School (-)	Can't Tell
(5, 6)	College (+)	High School (-)	Can't Tell

Part C (8 points): Party Time

All of Ben Bitbiddle's friends have ignored his seating plan. The friends' new seating locations are shown in the plot below.

For your convenience, a copy of the plot is provided on a tear-off sheet.



For each of the following kernels, determine whether Ben can use it to separate his friends (+ and **-**) into two groups.

Linear *(kernel of degree 1)*	**YES**	**NO**
Quadratic *(kernel of degree 2)*	**YES**	**NO**
Cubic *(kernel of degree 3)*	**YES**	**NO**
RBF *(radial basis function)*	**YES**	**NO**

Problem 3: Spiritual and Right Now (6 points)

For each question, **write in the box** provided the letter corresponding to the **one** best answer **and circle** the answer. There is **no penalty for wrong answers,** so it pays to guess in the absence of knowledge.

1. **AlphaGo** and **AlphaGo Zero** differ in that:
   A. AlphaGo uses tree search, and AlphaGo Zero does not.
   B. AlphaGo uses a single neural network, and AlphaGo Zero uses two neural networks working in sequence.
   C. AlphaGo performs lookahead, and AlphaGo Zero does not.
   D. AlphaGo was trained using data from games played by human experts, and AlphaGo Zero was not.

2. Sussman showed how **learning in sparse spaces** can be done with approximately:
   A. 2 examples.
   B. 20 examples.
   C. 200 examples.
   D. 2000 examples.

3. **Mutations** and **crossover** are useful in genetic algorithms because:
   A. Mutations enable spreading out while crossover enables hill-climbing.
   B. Mutations enable hill-climbing while crossover enables spreading out.
   C. Together they weed out unfit individuals from the population.
   D. Together they introduce all the necessary diversity in the population.

4. Defining a **transition vocabulary** is important in knowledge representation because:
   A. It captures the specificity of an event.
   B. It enables humans to represent and reason about events.
   C. It enables computers to represent and reason about events.
   D. It creates a hierarchy of classes.

5. According to Koile, Winston's goal in developing **near-miss learning** was to:
   A. Model how students learn with the help of a teacher.
   B. Demonstrate knowledge acquisition via inductive inference from examples.
   C. Make famous the idea of learning from examples.
   D. Interpret line drawings of scenes involving arches, tables, and tents.

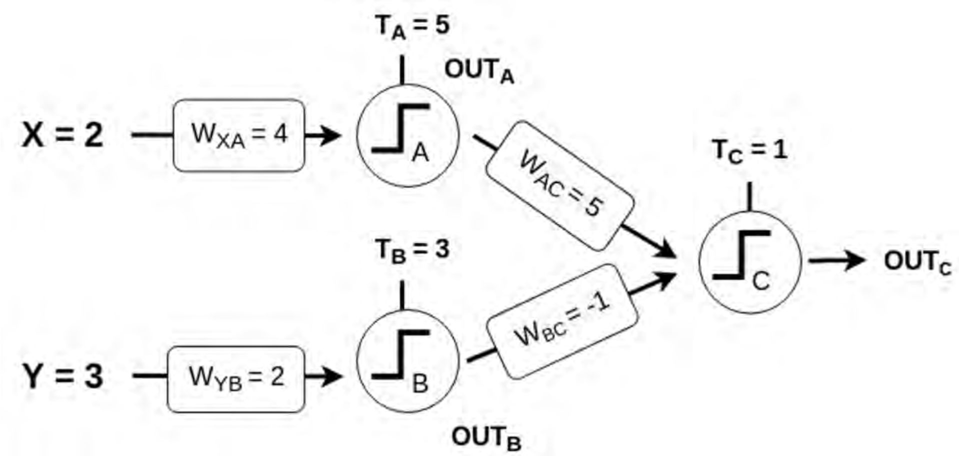
6. Kanwisher discussed using **convolutional neural networks** as evidence for her **hypothesis** that:
   A. Different parts of the brain are specialized to perform specific cognitive tasks.
   B. There are regions of the brain that do not have specialized functions.
   C. The ability to recognize faces is present in humans at birth rather than learned by experience.
   D. The ability to externalize language is not necessary for logical thought.

Tear-off sheet

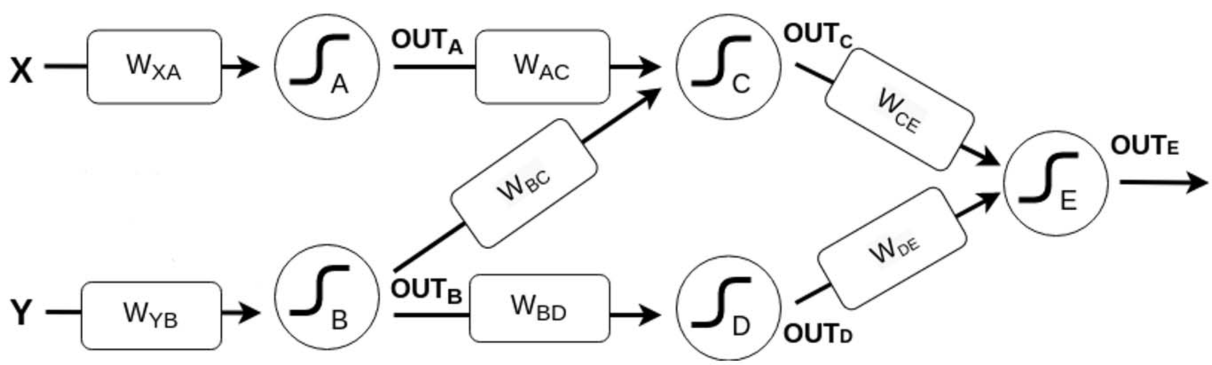
We do not collect tear-off sheets, so please show your work on the quiz pages, not the tear-off sheet.

Problem 1 (Neural Nets)

Part A2



Part B

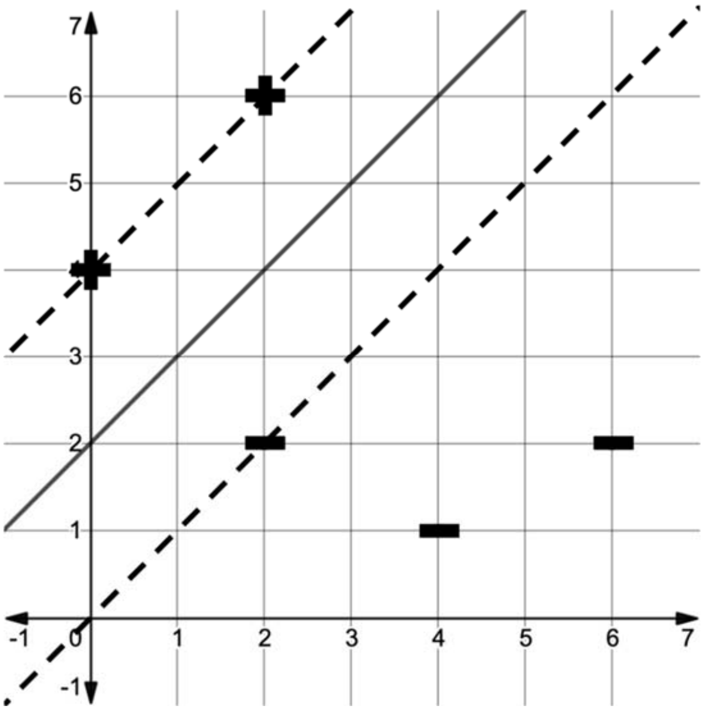


Tear-off sheet

We do not collect tear-off sheets, so please show your work on the quiz pages, not the tear-off sheet.

Problem 2 (SVM)

Part B



Part C