6.034 Quiz 3
5 November 2018

Name	SOLUTIONS
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.

Suri Bandler	**Marie Feng**	**Kifle Woldu**
Sanchit Bhattacharjee	**Ariel Jacobs**	**Matt Wu**
Alex Charidis	**Victoria Longe**	**Richard Yip**
Samir Dutta	**Smriti Pramanick**	

Problem number	Maximum	Score	Grader
1 - SVM	45		
2 – Neural Nets	55		
Total	100		

SRN	6		

There are 10 pages in this quiz, including this one, but not including tear-off sheets. Tear-off sheets with duplicate drawings are located after the final page of the quiz.

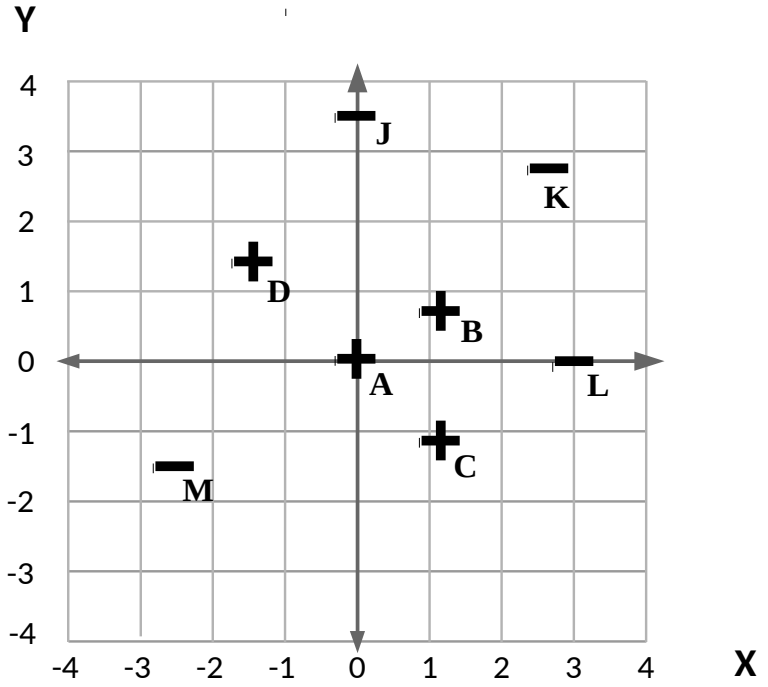
As always, open book, open notes, open just about everything, including a calculator, but no computers.

Problem 1: Support Vector Machines (45 points)

Part A: XOXO, SVM (45 points)

Enraged by Gossip Girl's most recent social media post, Blair and Chuck team up to expose Gossip Girl's secret identity. They determine that the suspects consist of people who attended the same party as Gossip Girl.

Blair and Chuck know of four **ATTENDEES** and four **NO-SHOWS**. The GPS locations of the four **ATTENDEES (+)** and four **NO-SHOWS (—)** are plotted below.



For your convenience, this graph is provided on a tear-off sheet at the end of the quiz.

A1 (5 points) Which of the following **kernel functions** can Blair and Chuck use to perfectly classify all people as **ATTENDEE** or **NO-SHOW** in the above dataset? Circle all that apply or circle NONE OF THESE if none apply.

LINEAR **QUADRATIC** **CUBIC** **RADIAL BASIS**

NONE OF THESE

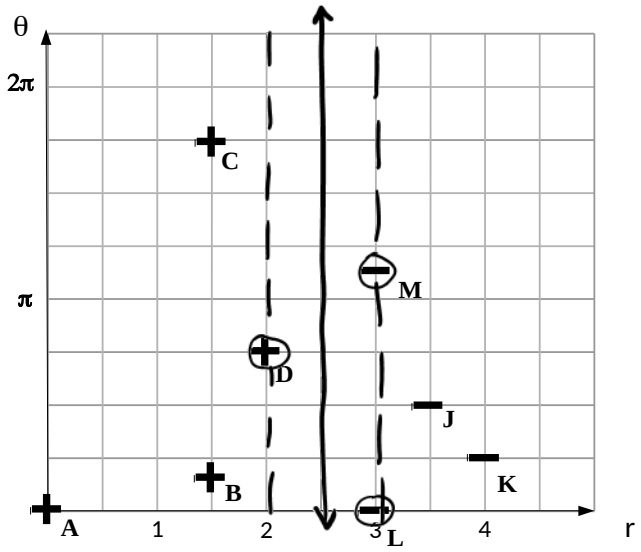
A2 (9 points) Blair and Chuck apply a polar transformation on the data, φ(<x, y>) = <r, θ >. Below is a graph of this transformation. *A duplicate of the graph is provided on the right and on a tear-off sheet.*

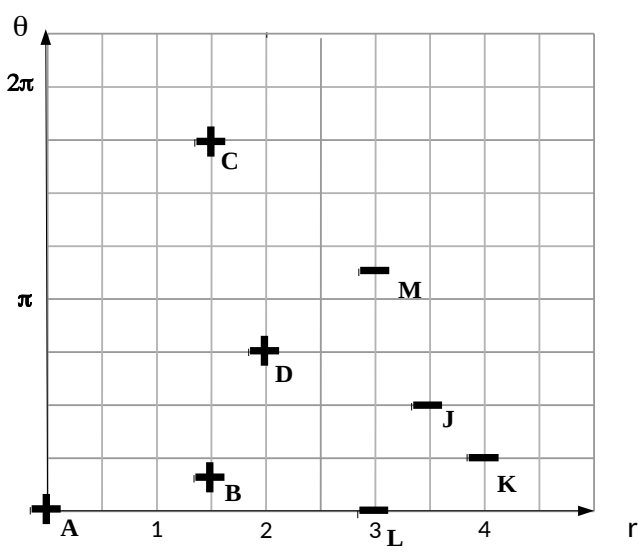
On the graph,

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

If you want us to grade the graph on the right, check the box.

This graph will be graded by default. I want to start over; grade this graph. ☐





A3 (8 points) For your decision boundary, calculate the values of the normal vector $\vec{w}$ and the offset ***b***, given the boundary equation $\vec{w}\cdot\vec{x} + b = 0$. The classifier must produce an output of **+1 for ATTENDEES (+)** and an output of **-1 for NO-SHOWS (–).**

$\vec{w} = [-2\ 0]$

$b = 5$

For partial credit for part A3, show your work here.

boundary $r = 2.5 \Rightarrow 1\cdot r + 0\cdot\theta - 2.5 = 0$

$[1\ 0]\begin{bmatrix} r \\ \theta \end{bmatrix} - 2.5 = 0$

$\vec{w}\cdot\vec{x} + b = 0$

$\vec{w} = k[1\ 0]$, $b = k(-2.5)$

find k.

method 1: margin width $= 1 \Rightarrow 1 = \frac{2}{\|\vec{w}\|} = \frac{2}{\sqrt{k^2+0^2}}$

$k = \pm 2$; because $\vec{w}$ points to positive $\Rightarrow k = -2$

method 2: choose a support vector + use $\vec{w}\cdot\vec{x} + b = \pm 1$

for $[3\ 0]$: $[k\ 0]\begin{bmatrix} 3 \\ 0 \end{bmatrix} + k(-2.5) = -1$

$k = -2$

$\therefore \vec{w} = [k\ 0] = [-2\ 0]$

$b = k(-2.5) = 5$

A4 (6 points) In each box below, list the points (A, B, C, D, J, K, L, M) for which the supportiveness (α) is equal to zero ($\alpha = 0$), greater than zero ($\alpha > 0$), and less than zero ($\alpha < 0$). If there are no such points, write NONE in the box. Each of the points should appear exactly once.

List all points for which:

… $\alpha = 0$ | A B C J K

… $\alpha > 0$ | D L M

… $\alpha < 0$ | None (d values are always ≥ 0)

A5 (5 points) Of the relations written below, circle **ALL** that are true. If none of the statements are true, instead circle NONE OF THESE. *Note: This problem can be solved without knowing the exact numbers for each* α.

$\alpha_M = \alpha_L$ $\alpha_M < \alpha_D$ $\alpha_B > \alpha_D$ $\alpha_L < \alpha_M$

NONE OF THESE

A6 (9 points) Blair and Chuck give you the GPS locations **in polar coordinates** of three people. For each of the three locations listed below, circle the ONE best answer indicating how the SVM would classify the person at that location.

($r = 1, \theta = \pi$):	ATTENDEE (+)	NO-SHOW (–)	CAN'T TELL
($r = 2.5, \theta = \pi/2$):	ATTENDEE (+)	NO-SHOW (–)	CAN'T TELL
($r = 3, \theta = 2\pi$):	ATTENDEE (+)	NO-SHOW (–)	CAN'T TELL

A7 (3 points) Consider your decision boundary found in polar coordinates. What does this boundary look like in the original, pre-transformed space? Circle the ONE best answer. If none of them are correct, circle NONE OF THESE instead.

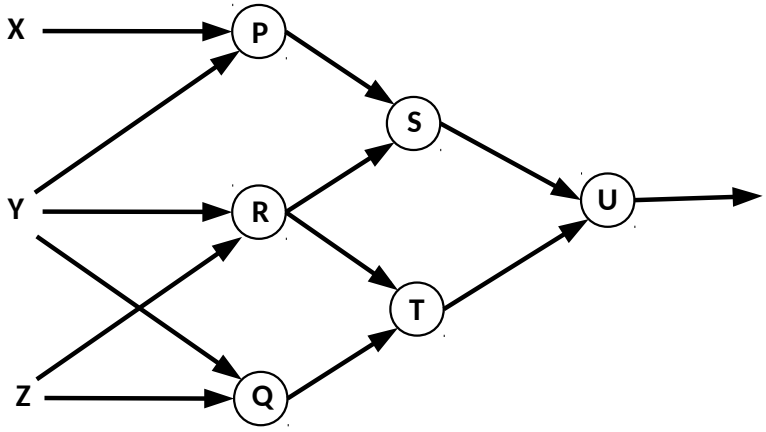
LINE **RECTANGLE** **TRIANGLE** **CIRCLE**

PARABOLA **NONE OF THESE**

Problem 2: Neural Networks (55 points)

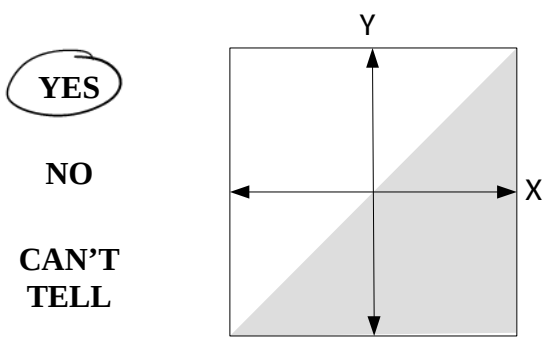
Part A: Funding Secured? (31 points)

Recently, the famed entrepreneur Melon Usk tweeted that his vehicle company secured a $421 billion buyout. The government devised a neural network to predict Usk's funding amount. The government has hired you, a famed 6.034 student, to analyze the network.



A1 (16 points) Assume the neural network shown above outputs 0 or 1 and that each neuron in the network uses a **stair-step activation function**. Each graph below depicts decision boundaries, with shaded regions corresponding to an output of 1.

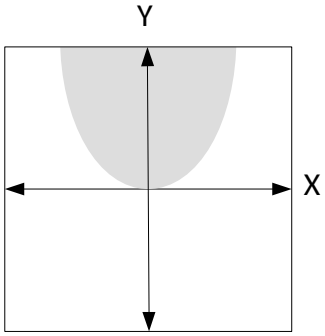
By changing thresholds and weights, which of the following graphs could the neural network depict? For **each** graph, circle YES, NO, or CAN'T TELL.



YES

NO

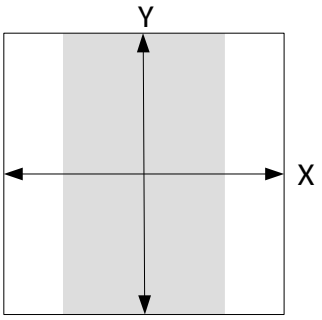
CAN'T TELL



YES

NO

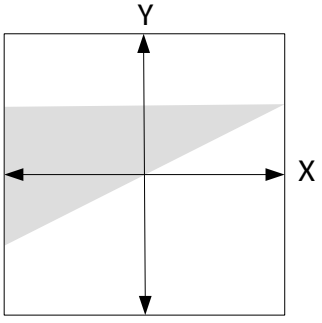
CAN'T TELL



YES

NO

CAN'T TELL



YES

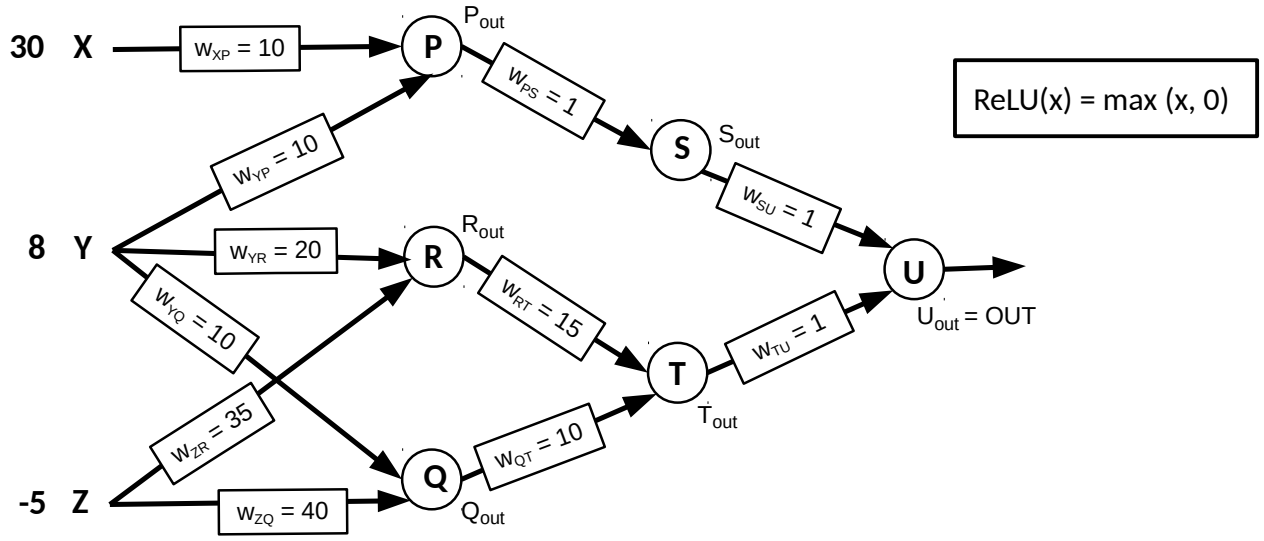
NO

CAN'T TELL

Note: The top boundary line in the above graph is horizontal.

A2 (10 points) Satisfied with your analysis of the network's architecture, the government now asks you to perform forward propagation on the network, using the weights provided below. You are given the initial inputs: **(X = 30, Y = 8, Z = - 5),** and you are told that each neuron uses the **ReLU activation function,** ReLU(x) = max (x, 0). Compute the output of each neuron and write the values in the table below. *For your convenience, this diagram is provided on a tear-off sheet at the end of the quiz.*

P_{out} = 380	S_{out} = 380	R_{out} = 0	T_{out} = 0	U_{out} = 380



For partial credit for part A2, show your work here.

$P_{out} = 30 \cdot 10 + 8 \cdot 10 = 380 > 0 \Rightarrow P_{out} = 380$

$R_{out} = 8 \cdot 20 + (-5) \cdot 35 = -15 < 0 \Rightarrow R_{out} = 0$

$Q_{out} = 8 \cdot 10 + (-5) \cdot 40 = -120 < 0 \Rightarrow Q_{out} = 0$

$T_{out} = R_{out} \cdot 15 + Q_{out} \cdot 10 = 0 \Rightarrow T_{out} = 0$

$S_{out} = P_{out} \cdot 1 = 380 > 0 \Rightarrow S_{out} = 380$

$U_{out} = S_{out} + T_{out} = 380 > 0 \Rightarrow U_{out} = 380$

A3 (1 point) The final output of the network, U_{out}, corresponds to Usk's predicted funding amount, in billions of dollars. Did the network predict Usk's funding to be at least 421?

Circle one.

YES **NO** **CAN'T TELL**

A4 (4 point) Usk's friend Beff Jezos argues that it's possible to change the value of **one** of the weights in the network in order to produce an output of at least 421.

To which weight in the list below is Jezos referring? Circle **one** weight and write its corresponding new value in the box below. Each weight is named by the two nodes that connect it. If it's not possible to change one of these weights to get an output of at least 421, circle NONE. **Assume all weights only have integer values.**

Weight: W_{RT} W_{ZQ} OR W_{YP} W_{TU} NONE **New Value:**

$U_{out} = S_{out} + T_{out} \geq 421$

(1) increase S_{out} by increasing P_{out}

$30 \cdot 10 + W_{YP} \cdot 8 \geq 421$

$W_{YP} \geq 16$ for integer values only

OR (2) increase T_{out} by increasing Q_{out}

$T_{out} \geq 421 - 380 \geq 41$

$T_{out} = R_{out} \cdot 15 + Q_{out} \cdot 10$; $Q_{out} \cdot 10 \geq 41$; $Q_{out} \geq 4.1$

$\therefore Q_{out} \geq 5$ because integer values only

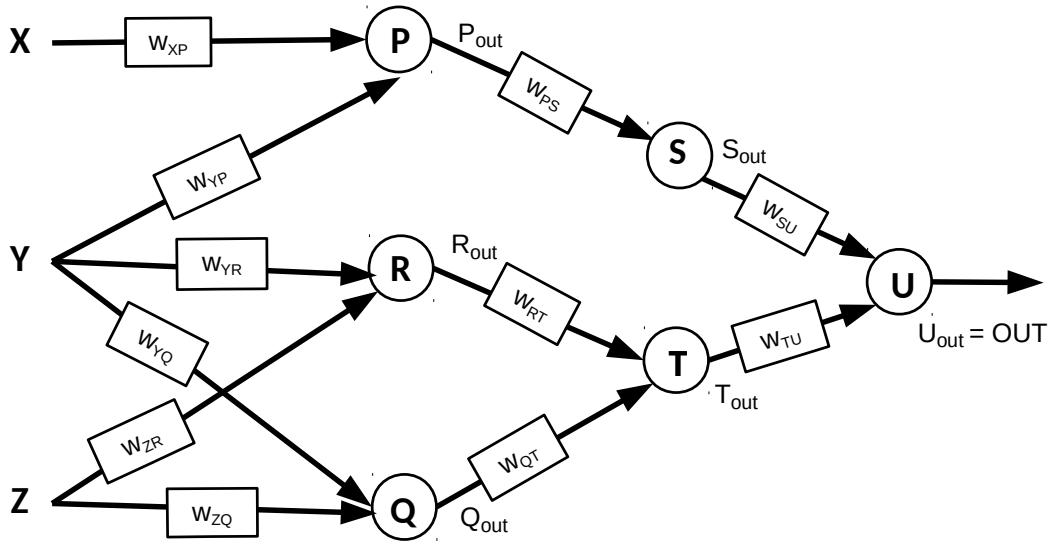
OVER FOR PART B →

$Q_{out} = 8 \cdot 10 + (-5) W_{ZQ} \geq 5$

$W_{ZQ} \leq 15$ for integer values only

Part B: Backpropagation (24 points)

When questioned, Usk claims that the government's network diagram models revenues for his other company, SolarCity. He claims that the network diagrammed below is correct, but the government is suspicious of the network's dependencies.



For your convenience, this diagram is provided on a tear-off sheet at the end of the quiz.

B1 (15 points) Labels for each input, weight, and neuron output are shown on the network above. For example, R_{out} is the output for neuron R. Which of the following choices is an expression for the calculation of ΔW_{PS} given the labels in the figure above?

Assume that this network uses a **sigmoid threshold function,** has a **learning rate of 2**, and has a **desired output of 0.5**. Circle the one best answer.

For partial credit, there is space on the next page for you to show your work.

$\Delta W_{PS} =$

A. $(2)(-\frac{1}{2})(0.5 - U_{out})^2$

B. $2P_{out}(S_{out}(1 - S_{out})(W_{XP}(x\ (1-x)(0.5-x)) + W_{YP}(y\ (1-y)(0.5-y))))$

C. $2P_{out}(S_{out}(1 - S_{out})W_{SU})$

D. $2P_{out}(S_{out}(1 - S_{out})W_{SU}(U_{out}(1 - U_{out})(0.5 - U_{out})))$

E. $2P_{out}S_{out}(1 - S_{out})(W_{SU} + W_{TU})(0.5 - U_{out})$

For partial credit for part B1, show your work here.

$\Delta w_{ps} = r \cdot P_{out} \cdot \delta_s$

$\delta_s = s_{out}(1 - s_{out}) w_{su} \delta_u$

$\delta_u = u_{out}(1 - u_{out})(0.5 - u_{out})$

$\Delta w_{ps} = 2 \cdot P_{out} \cdot s_{out}(1 - s_{out}) w_{su} \cdot u_{out}(1 - u_{out})(0.5 - u_{out})$

B2 (9 points) The government likes your work on ΔW_{PS} and now wants your advice about the dependencies in backpropagation for a different weight, W_{XP}.

W_{XP} depends on which <u>weights</u>? List the weights' labels in the box below.

W_{PS} W_{SU} (W_{XP} also accepted)

W_{XP} depends on which neuron <u>inputs</u>? List the neuron inputs' labels in the box below.

X

W_{XP} depends on which neuron <u>outputs</u>? List the neurons' labels in the box below.

P_{out} s_{out} u_{out}

OVER FOR SPIRITUAL AND RIGHT NOW QUESTIONS ⟶

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.

[c] 1. Winston claimed that the best way to learn the difference between the sounds *bah* and *gah*, or more generally, to learn how to pronounce sounds in a new language is to:

a. Listen to audiotapes, because you can easily and repeatedly listen to specific pronunciations.
b. Read books in the language so that you can see how words are spelled.
c. Watch a native speaker speak so that you can see how they move their lips.
d. None of the above; there is no good way to learn as an adult.

[b] 2. Marr's approach to problem solving emphasized:

a. Applying successful methods to as wide of a variety of problems as possible for efficient and prolific results.
b. Understanding the problem in order to develop a method that best fits the problem you are trying to solve.
c. Ignoring implementation so the solution remains as general and far-reaching as possible.
d. User studies, as they are the only way to determine if your solution makes sense in the real world.

[d] 3. Yuret argued that a neural net understands what it sees in an image if it can:

a. Generate a new picture that has similar properties (content, texture, etc.).
b. Distinguish between classes of objects, such as dogs and cats.
c. Generate a caption that describes the image.
d. Answer questions about the image.

[b] 4. Koile argued that Karl Sims's use of genetic algorithms to evolve creatures was so successful because:

a. The project was well funded by a video game company, due to the project's focus on computer graphics.
b. The domain had many solutions, allowing mutation and crossover operations to produce a variety of outcomes.
c. He used a simple fitness measure based on the number and size of objects.
d. He skillfully used deep neural networks to predict which mutations and genetic crossovers would be most effective for the next generation.

[c] 5. Winston's near miss learning specializes knowledge through:

a. Discarding contradictory samples.
b. Creating negative examples to extend the set of training data.
c. Forbidding and requiring relation links.
d. Descending a hierarchical tree of object types.

[a] 6. Maes described closed loop enhancement of technology as:

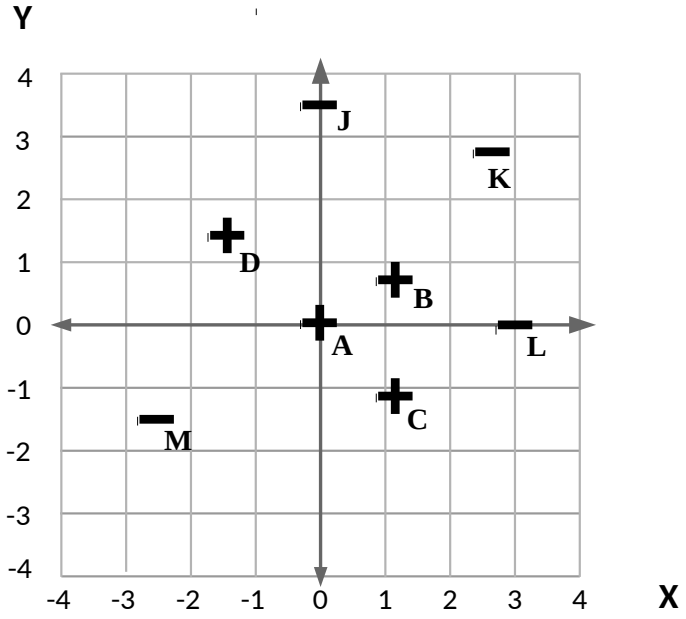
a. Real time sensing and modeling, followed by real time intervention and feedback.
b. Real time sensing, followed by post-processing, feedback, and enhanced sensing.
c. Crowdsourcing data, followed by creation of a generalizable model.
d. Real time probing for user input, followed by creation of user-specific models.

Tear-off sheet

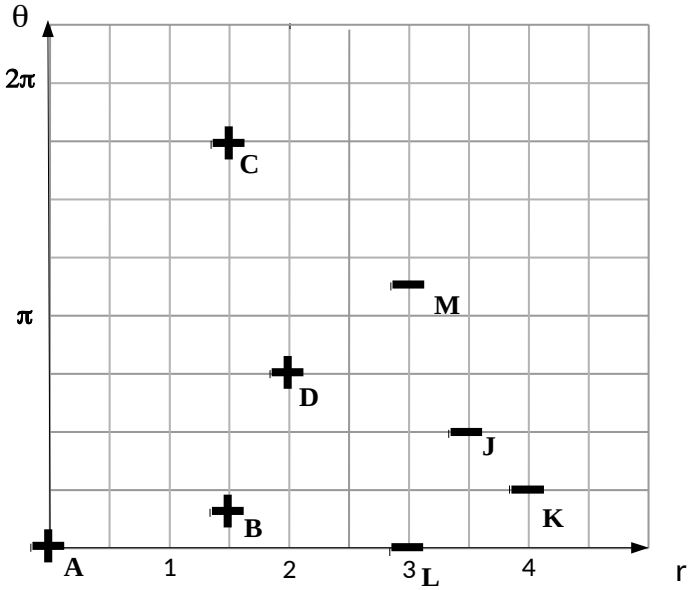
We do not collect tear-off sheets, so please show your work on the quiz pages, not the tear-off sheet.

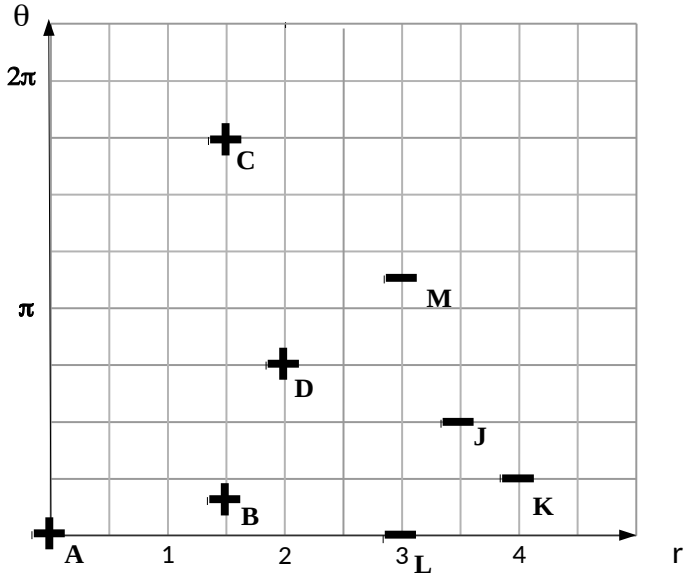
Problem 1 SVM

Part A1



Part A2



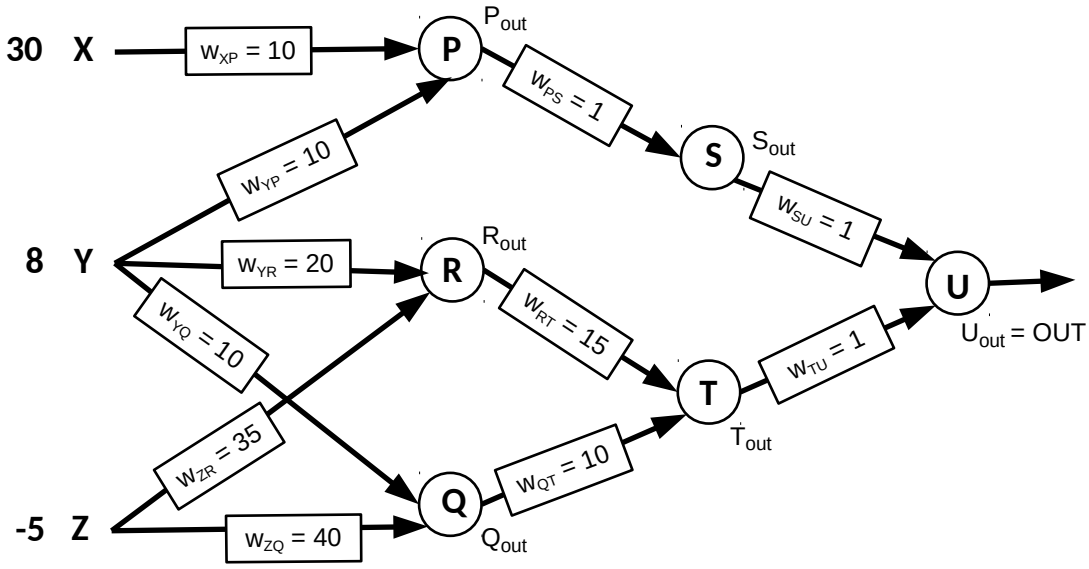


Tear-off sheet

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

Problem 2 Neural Networks

Part A2



Part B