

Problem Set 1
Due date: 11/14/18
Total Points: 40

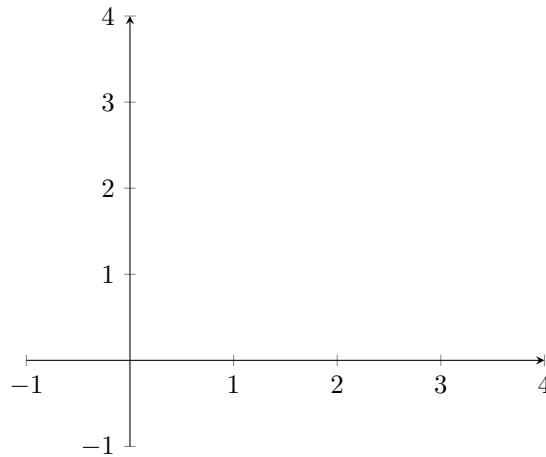
If you run out of room for an answer, use scratch paper and staple it to this sheet.

Name and Grade: _____

1. (1 point) Given the following data:

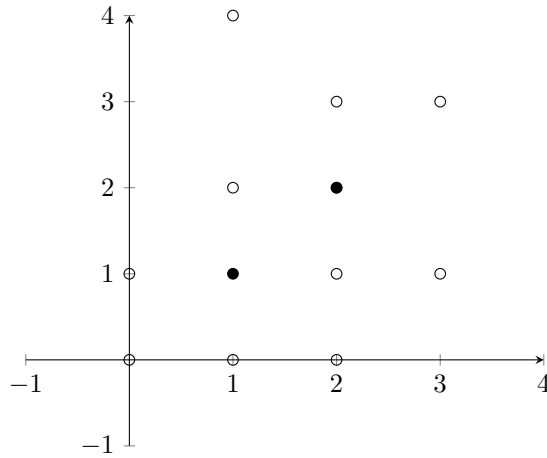
Graph the data with the horizontal axis being x_1 and the vertical axis x_2 . Shade in points that correspond to $y = 0$ and leave hollow the ones that correspond to $y = 1$.

| x_1 | x_2 | y |
|-------|-------|-----|
| 0 | 2 | 0 |
| 1 | 1 | 0 |
| 1 | 2 | 0 |
| 1 | 4 | 1 |
| 2 | 1 | 0 |
| 3 | 2 | 1 |
| 3 | 3 | 1 |



2. (1 point) Draw the decision boundary (line) created by a single perceptron that classifies the data perfectly.
3. (2 points) Write the equation for the decision boundary created above assuming the perceptron outputs 1 when $f(x) > 0$ and 0 when $f(x) < 0$.
4. (8 points) You are now given a perceptron with a decision boundary $x_2 + x_1 - 4 = 0$. Assume the point at (3,2) is moved to (3,1). Use learning rate $\alpha = 0.2$.
- Graph this new decision boundary.
 - Update the decision boundary using the misclassified point.
 - Was α too high, too low, or fine? (circle one)
 - If α was too high or too low, find a value of α that leads to perfect classification after one update step.

5. (10 points) You are given the following data. Hollow points correspond to class 0, shaded points correspond to class 1.



- (a) Draw a Multi-Layer Perceptron that perfectly classifies the data above. Make sure to write out the values for every weight and bias in your MLP.
- (b) On the graph above, draw the decision boundaries that correspond to each perceptron in your MLP.
6. (10 points) You have a neural network with three layers of nodes. Assume all input values are 1 and all biases are -1. Use the activation function $f(x) = x^2$. Draw a diagram of the network and write the output for each node in the final layer.

$$W_1 = \begin{bmatrix} 3 & 3 & 0 & 1 \\ 1 & 2 & 2 & 3 \\ 3 & 4 & 1 & 5 \\ 1 & 3 & 0 & 2 \end{bmatrix} \quad W_2 = \begin{bmatrix} 1 & 2 & 2 & 0 \\ 4 & 4 & 2 & 2 \\ 2 & 0 & 1 & 1 \end{bmatrix}$$

7. (3 points) The neural network learning algorithm (backpropagation) will be covered next week. Backpropagation requires a differentiable activation function, which is why neural networks do not use the Perceptron's step function. A popular activation function is the sigmoid function.

$$\sigma(x) = \frac{1}{1 + e^{-x}}$$

- (a) The sigmoid function is differentiable. Find its derivative.

- (b) Rewrite $\sigma'(x)$ in terms of $\sigma(x)$.

8. (2 points) What are the advantages of the Sigmoid function, $\sigma(x) = \frac{1}{1 + e^{-x}}$, over the step function? Circle all that apply.

- (a) Sigmoids are differentiable.
 (b) $\sigma(x)$ quickly approaches 1 as $x \rightarrow \infty$ and $\sigma(x)$ quickly approaches -1 as $x \rightarrow -\infty$, so networks learn quickly.
 (c) Sigmoids have a simple derivative in terms of the original function, so they are faster to compute than a step function.
 (d) The inflection point of $\sigma(x)$ is located on the y-axis, which leads to faster computations with vectorized representations
 (e) Sigmoids can tell you how far away you are from the correct value based on the magnitude of x

9. (3 points) There are 16 two-valued boolean functions.

| A | B | F ₀ | F ₁ | F ₂ | F ₃ | F ₄ | F ₅ | F ₆ | F ₇ | F ₈ | F ₉ | F ₁₀ | F ₁₁ | F ₁₂ | F ₁₃ | F ₁₄ | F ₁₅ |
|---|---|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 0 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 |
| 1 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 1 |
| 1 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 1 |

- (a) Which of the boolean functions above cannot be perfectly classified by a single perceptron?

- (b) What is/are the name(s) of the function(s) you found in part (a)? (A AND B, A NOT B, etc.)