1. (True/False) A linear classifier can only learn positive coefficients.                      1 point

   ○ True
   ● False


2. (True/False) In order to train a logistic regression model, we find the weights that          1 point
   maximize the likelihood of the model.

   ● True
   ○ False


3. (True/False) The data likelihood is the product of the probability of the inputs $\mathbf{x}$ given    1 point
   the weights $\mathbf{w}$ and response $y$.

   ○ True
   ● False


4. Questions 4 and 5 refer to the following scenario.                                             1 point


   Consider the setting where our inputs are 1-dimensional. We have data

   | $x$ | $y$ |
   |-----|-----|
   | 2.5 | +1 |
   | 0.3 | -1 |
   | 2.8 | +1 |
   | 0.5 | +1 |


   and the current estimates of the weights are $w_0 = 0$ and $w_1 = 1$. ($w_0$: the intercept, $w_1$:
   the weight for $x$).
   Calculate the likelihood of this data. Round your answer to 2 decimal places.

   0.23

5. Refer to the scenario given in Question 4 to answer the following:

1 point

Calculate the derivative of the log likelihood with respect to $w_1$. Round your answer to 2 decimal places.

0.37

6.

1 point

Which of the following is true about gradient ascent? Select all that apply.

- ☑ It is an iterative algorithm
- ☐ It only updates a few of the parameters, not all of them
- ☑ It finds the maximum by "hill climbing"