Chapter 10: Theoretical Questions

- Show that the MLP in Figure 10.6 of the book solves the XOR problem.
- (Exercise 5 from the book) Name and draw three popular activation functions.
- (Exercise 6 from the book) Suppose you have an MLP composed of one input layer with 10 passthrough neurons, followed by one hidden layer with 50 artificial neurons, and finally one output layer with 3 artificial neurons. All artificial neurons use the ReLU activation function.
  - What is the shape of the input matrix $\mathbf{X}$?
  - What are the shapes of the hidden layer's weight matrix $\mathbf{W_h}$ and bias vector $\mathbf{b_h}$?
  - What are the shapes of the output layer's weight matrix $\mathbf{W_o}$ and bias vector $\mathbf{b_o}$?
  - What is the shape of the network's output matrix $\mathbf{Y}$?
  - Write the equation that computes the network's output matrix $\mathbf{Y}$ as a function of $\mathbf{X}$, $\mathbf{W_h}$, $\mathbf{b_h}$, $\mathbf{W_o}$ and $\mathbf{b_o}$.
- (Exercise 7 from the book) How many neurons do you need in the output layer if you want to classify email into spam or ham? What activation function should you use in the output layer? If instead you want to tackle MNIST, how many neurons do you need in the output layer, and which activation function should you use? What about for getting your network to predict housing prices, as in Chapter 2?
- Explain why activation functions are necessary in neural networks.
- Suppose the logits are $(-1, 0, 2)$ for a classification task with three classes. What are the probabilities for each class if we use the softmax activation function?
- Describe how to construct a neural network that is equivalent to a logistic regression model.
  - How many layers does it need?
  - What activation function should you use in the output layer?
  - What loss function should you use?
- Consider a TLU with weights $w_1 = 2$, $w_2 = -1$ and bias $b = 1/2$.
  - Is the example $(x_1, x_2) = (1, 1)$ classified as positive or negative?
  - What about the example $(x_1, x_2) = (-1, 0)$?
  - In the (x_1, x_2)-plan, sketch the decision boundary of this TLU and indicate which side is the positive class.