1) **A = TRUE**

   Why? Slide 2 on Dr. Race's notes→ Neural networks have the potential to model very complicated patterns and can be used for both classification and continuous prediction tasks. Hence, they are "universal approximators."

   B = TRUE

   Why? Slide 21 on Dr. Race's notes → Neural networks have no mechanisms for variable selection. You provide inputs and it is simply trying to model the relation between the variables.

   C = FALSE

   Why? Slide 21 on Dr. Race's notes → Neural networks result in very complicated relationships so they are hard to interpret.

   D = FALSE

   Why? The advantage with neural networks is that they can model nonlinear relationships. Remember Dr. Race's example of neural networks being able to create a mountainous surface model → Slide 19 on Dr. Race's notes enforced this idea through an example.

2) Here is the example from class from slide 8:
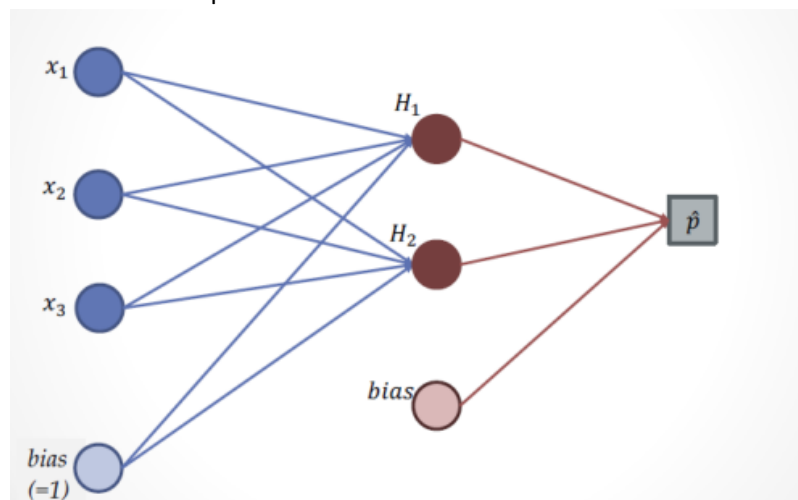
   Using the above diagram, the solution to the problem in the exercise is:

   1 parameter associated with going from input 1 to each of the 5 hidden units → 1*5 = 5

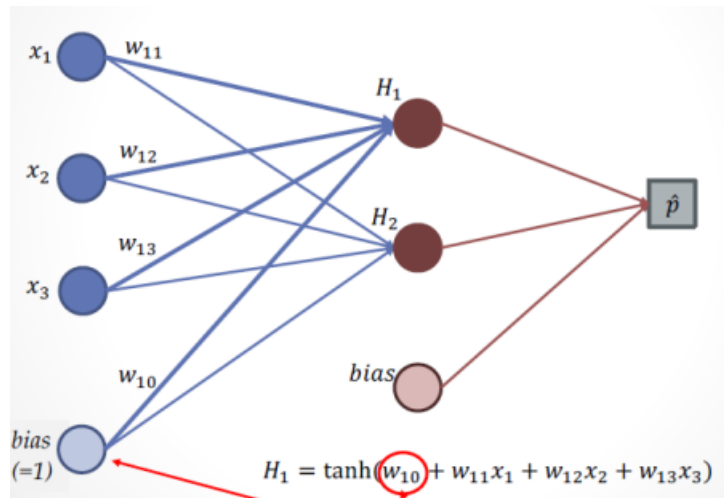   1 parameter associated with going from input 2 to each of the 5 hidden units → 1*5 = 5

   1 bias from the input layer going to each of the 5 hidden units → 1*5 = 5

   1 parameter representing going from each hidden layer to output → 1*5 = 5

   1 bias from the hidden layer to the output → 1

   Total = 5+5+5+5+1 = 21 parameters needed

3) Slide 9 from Dr. Race's notes:

Based on this slide, the answer to the exercise is:

3 is the bias associated with going from the input layer to the hidden unit (the intercept term). However, based on the information provided, we cannot determine the bias associated with going from the hidden layer to the output because we aren't given the formula.

4) **TRUE** → Slide 13 of Dr. Race's notes

5) **FALSE** → Slides 9 – 10 of Dr. Race's notes; Logistic regression goes from 0 to 1 while the hyperbolic tangent function goes from -1 to 1.

6) Using slides 11 – 13 to serve as reference:

Slide 11: Going from inputs to hidden unit 1
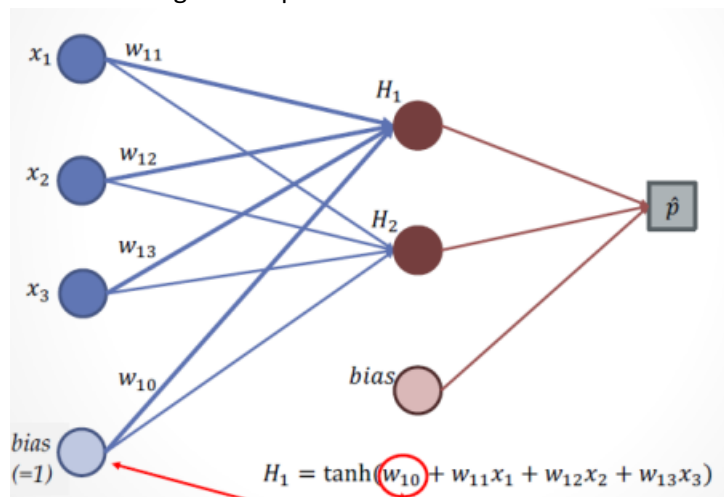
Based on the above figure:

w10 = BIAS_H11 = -1.117641

w11 = fineagg_H11 = 2.049942

w12 = slag_H11 = 0.605445

w13 = superplastic_H11 = -2.294231

Slide 12: Going from inputs to hidden unit 2

$$H_2 = \tanh(w_{20} + w_{21}x_1 + w_{22}x_2 + w_{23}x_3)$$
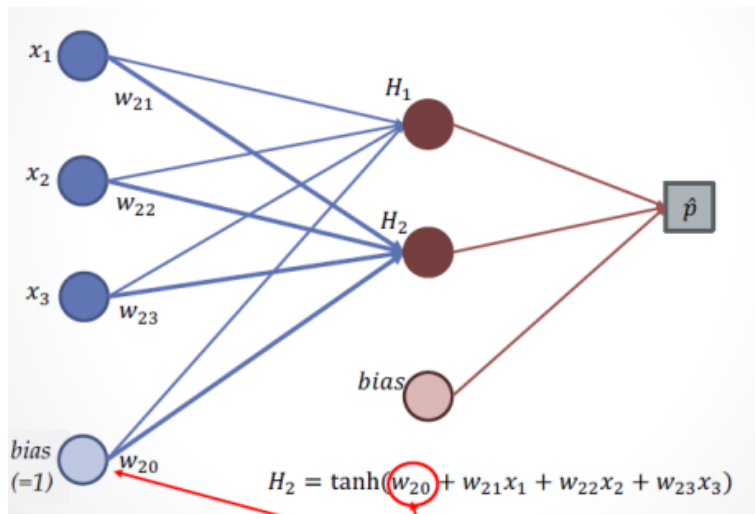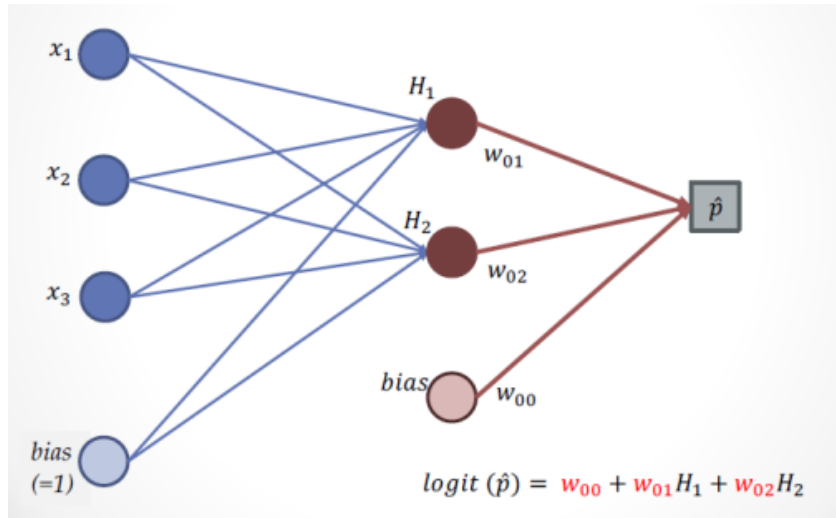
Based on the above figure:

w20 = Bias_H12 = 3.832763

w21 = fineagg_H12 = -1.308505

w22 = slag_H12 = -2.626001

w23 = superplastic_H12 = -2.509042

Slide 13: Going from hidden layer to output

$$\text{logit}\,(\hat{p}) = w_{00} + w_{01}H_1 + w_{02}H_2$$

Based on the above figure:

w00 = Bias_strength = 38.415742

w01 = H11_strength = -7.739118

w02 = H12_strength = -7.258123