

1. Support vector machines always find linear decision boundaries in the original feature space.

- ☐ A True
- ☐ B False

2. Support vectors are the positives and negatives used to form the decision boundary.

- ☐ A True
- ☐ B False

3. Support vector machines minimize the margin between positive and negative samples.

- ☐ A True
- ☐ B False

4. Support vector machines minimize the magnitude of the weight vector.

- ☐ A True
- ☐ B False

5. What is the kernel trick?

- ☐ A Lifting points to a higher-dimensional space
- ☐ B Not having to define the lifting function explicitly for a single argument
- ☐ C Letting the method figure out the right lifting transformation

6. Support vectors have non-zero alpha weights.

- ☐ A True
- ☐ B False

7. For most neural networks, the outputs are a linear combination of the inputs (i.e.  $y = W * x$ ).

- ☐ A True
- ☐ B False

8. We have a network with 3 inputs (no bias) and 2 hidden neurons in a single hidden layer.  $w(1)_{11} = 0.1$ ,  $w(1)_{12} = 0.2$ ,  $w(1)_{13} = 0.7$ ,  $x_1 = 10$ ,  $x_2 = 10$ ,  $x_3 = 0$ . At the hidden layer, we have  $h(a) = a^2$  (a-squared). What is  $z_1$  equal to?

- ☐ A 1
- ☐ B 3
- ☐ C 9
- ☐ D 20

9. To train a network, we compute a set of quantities  $\delta$ , which we use to update the weights. For which nodes do we compute  $\delta$  first?

- ☐ A inputs
- ☐ B hidden
- ☐ C outputs

10. The loss is a function of the activations at the hidden layer, which are a function of the weights, so we need to compute:  $dE/dz_j * dz_j/daj * daj/dw_{ji}$

- ☐ A True
- ☐ B False

11. We might use backprop of error to compute the value for which of the above quantities?

- ☐ A  $dE/dz_j$
- ☐ B  $dz_j/daj$
- ☐ C  $daj/dw_{ji}$

12. What is the key intuition for using ensembles?

- ☐ A the imperfect individual learners will make uncorrelated mistakes
- ☐ B we will only include very strong individual learners
- ☐ C individual learners will rely on each others' answers to collectively perform better